Prognosis of cognitive change following stroke: Identifying predictors, investigating the nature of between-variable relationships, and accounting for individual variability

A Thesis by

Bogna Anna Drozdowska (MSc)

Submitted for the degree of Doctor of Philosophy

To

The University of Glasgow

From

The Institute of Cardiovascular and Medical Sciences

College of Medical, Veterinary and Life Sciences

University of Glasgow

Glasgow Royal Infirmary

2021

Abstract

Background: Stroke survivors are at high risk of experiencing cognitive problems, which can severely compromise independence in daily activities, social participation, and quality of life. With limited evidence to support current interventions, finding ways to improve cognitive function is recognised as a priority for research relating to life after stroke. Prognosis research may contribute to this endeavour by informing the development and implementation of preventive and therapeutic strategies - from describing the natural history of post-stroke cognitive change, through identifying its relevant predictors and developing methods of estimating individual outcome probability, to supporting the application of stratified medicine.

Prognosis research into post-stroke cognition is still developing, with little evidence regarding some of its more fundamental questions. These relate to the relevance of: i) potentially modifiable factors, ii) differential effects of risk factors, depending on paths of influence and co-occurrence, and iii) population heterogeneity in the trajectory of post-stroke cognitive change. Through focusing on these three topics, the purpose of this thesis is to improve our understanding of the cognitive change that occurs following stroke and its associations with individual characteristics.

Methods: Firstly, to gain a better insight into current advances in prognosis research in post-stroke cognition, I performed a systematic review of prognostic rules for predicting cognitive impairment and delirium following stroke. I considered these findings in specifying the aims and design of my subsequent, observational studies.

I conducted two cross-sectional investigations in a sample of stroke survivors from the UK Biobank. Through a series of regression analyses, I assessed the associations of performance on four cognitive tasks with two groups of predictors of particular interest: 1) self-reported physical activity and sedentary behaviour, and 2) proxies of social engagement.

Using data from consecutive patients admitted to a hyper-acute stroke unit, I then investigated the influence of cardiovascular risk factors on acute post-

stroke cognitive performance. In a moderated mediation analysis, I tested the assumptions that the effects of these factors are partially mediated by stroke severity and prior dementia, and may be dependent on comorbidity.

In my final, longitudinal study, based on the Assessing Post-Stroke Psychology Longitudinal Evaluation (APPLE) dataset, I conducted a latent class growth analysis to identify and describe differential trajectories of cognitive change, occurring over one year following stroke. Through subsequent regression analyses, I then explored factors that predicted trajectory class membership.

Findings: Through a systematic review of the literature, I identified seven prognostic rules predicting post-stroke cognitive impairment (including dementia) and four predicting post-stroke delirium. The most commonly incorporated predictors were: demographics, imaging findings, stroke type, and symptom severity. Among seven studies that assessed in the original sample how well a prognostic rule discriminated between participants who developed the outcome of interest and those who did not, performance was reported as being good to excellent. Only one rule had been validated in an independent dataset, showing fair discriminatory power.

In the first of two UK Biobank studies, I found relatively consistent, although weak associations for two types of sedentary behaviour, where the daily duration of watching TV was associated with poorer cognitive performance, while duration of computer use was associated with better performance. Some effects remained significant after adjusting for demographic, health-related, and lifestyle factors. Physical activity, however, was not independently associated with performance on any of the considered tasks. In the second study, reported loneliness was the only proxy of social engagement to be associated with most cognitive tasks, consistently predicting poorer performance.

Findings from my analysis of data from a hyper-acute stroke unit setting supported the mediatory role of stroke severity and prior cognitive impairment in the effects of specific cardiovascular risk factors on acute cognition. Poorer cognitive performance was associated with atrial fibrillation through increased stroke severity, and with previous stroke through an increased risk of prevalent

dementia. Conversely, through an association with reduced stroke severity, better performance seemed predicted by vascular disease (in the presence of hypertension and absence of diabetes) and by previous transient ischaemic attack.

In the APPLE dataset, I identified four distinct trajectories of cognitive change: i) with high early cognitive function, improving over following weeks and thereafter declining; ii) with some early cognitive deficits, followed by improvement in function and then relative stability; iii) with comparatively poor initial function, which after a stage of steeper improvement continued to improve at a slower rate; and iv) with severe cognitive deficits, followed by improvement at a near-constant rate. Overall, participants representing the two trajectories with greatest initial cognitive deficits were characterised by older age, lower education, higher prevalence of pre-stroke cognitive impairment, and greater stroke severity.

Conclusions: In summary, my findings speak to the complex nature of cognitive change following stroke and its associations with individual characteristics. This is apparent on more than one level. What can be considered a single variable, such as sedentary behaviour, may be multifaceted. Entailing distinct properties, particular variable components are likely to have differential effects on post-stroke cognitive function. The effects of specific factors may moreover differ depending on the path of influence and the constellation of coexisting variables. Finally, post-stroke cognitive change is a heterogenous process, both on a between- and within-individual level.

These observations suggest that it is important to consider how, in what form, under what conditions, and for whom, a possibly causal factor can affect post-stroke cognitive outcome. A lack of evidence-based assumptions regarding these aspects to inform the development of a statistical model may lead to misidentification of relevant associations. This is in turn likely to have implications at the stage of intervention development and implementation, limiting application. Recognising and at least partly accounting for the complexities I observed in my series of studies could contribute to bridging a gap between the potential and actual impact of prognosis research on improving cognitive function following stroke.

Table of Contents

List of Tables

List of Figures

Acknowledgement

First and foremost, I would like to thank my supervisors, Dr Terence Quinn and Professor Peter Langhorne. I wish I had the eloquence to fully express my gratitude for their guidance, encouragement, and kindness, and for setting such an inspiring example for me. Despite the lapse of four years, I can still hardly believe how fortunate I was to be welcomed to their team as a PhD student.

I am grateful to the Stroke Association for their funding, without which conducting this work would not have been possible.

I wish to thank my colleagues at the New Lister for creating a truly supportive and friendly environment, which I will continue to treasure regardless of where I end up in the future. I am particularly appreciative of the time I spent with my fellow PhD student and dear pal, Emma Elliott. I have shared the ups and downs of this experience with her with such pleasure that even the latter of the two seemed amusing.

I would like to express my gratitude to those who made working on the APPLE project (a huge part of my PhD) so enjoyable. This includes the APPLE participants, to whom I owe special thanks – not just for their absolutely crucial involvement in the study, but also for making me look forward to each follow-up visit. I also wish to acknowledge Ruth Graham, for her master recruitment skills and for lighting up our office whenever she popped in.

I am deeply and enduringly grateful to my parents and sister, who put the wind in my sails in the first place, and continue making sure I never go without it.

Finally, I would like to begin thanking my fiancé, Danny, who put his own aspirations on hold so that I could fulfil mine. I now find myself in an extremely unique position, where the prospect of long-term debt repayment is a cause for immeasurable joy.

Author's Declaration

My contribution to each thesis chapter, together with significant contributions of others who were involved in conducted projects, is outlined below.

Dr Terence Quinn and Professor Peter Langhorne provided supervision, guidance, and feedback for all chapters of this thesis.

Chapters 1, 2, 6, and 9: I wrote these chapters.

Chapter 3: A previous version of the review was submitted to the University of Glasgow for the degree of Bachelor of Science by Michael McKay, who together with Dr Terence Quinn conceived the idea for the project. In consultation with my supervisors and a Cochrane Information Specialist, Joshua Cheyne, I redesigned the search strategy and repeated the literature search. Together with Dr Kris McGill, we screened retrieved publications, and carried out the data extraction and quality assessment for included studies. I wrote the chapter.

Chapter 4: I conceived the idea for both studies. Dr Carlos Celis-Morales and Dr Donald Lyall assisted me in using the UK Biobank database and in planning the analyses, including the development of analysis scripts for use in Stata software. I conducted the analyses and wrote the chapter.

Chapter 5: I contributed to data collection for the used database alongside other researchers: Robert Shaw, Martin Taylor-Rowan, Emma Elliott, and Gillian Cuthbertson. I conceived the idea for the study, planned and conducted the analyses, and wrote the chapter.

Chapters 7 and 8: I recruited participants and collected data for the APPLE project alongside two other PhD students (Emma Elliott and Martin Taylor-Rowan) and research nurses at each hospital site. Dr Terence Quinn and I conceived the idea for the study. I planned and conducted the analyses, and wrote the chapters.

Publications and Conferences

Publications related to this thesis

Chapter 3: Drozdowska, B. A., McGill, K., McKay, M., Bartlam, R., Langhorne, P., & Quinn, T. J. (2021). Prognostic rules for predicting cognitive syndromes following stroke: A systematic review. *European Stroke Journal*, 6(1), 18-27.

Chapter 4: Drozdowska, B. A., Celis-Morales C. A., Lyall, D. M., & Quinn, T. J. (2019). Social engagement after stroke – is it relevant to cognitive function? A cross-sectional analysis of UK Biobank data. *AMRC Open Research*, 1:3.

Chapter 5: Drozdowska, B. A., Elliott, E., Taylor-Rowan, M., Shaw, R. C., Cuthbertson, G., Langhorne, P., & Quinn, T. J. (2020). Cardiovascular risk factors indirectly affect acute post-stroke cognition through stroke severity and prior cognitive impairment: a moderated mediation analysis. *Alzheimer's Research & Therapy*, 12(1), 1-10.

Publications not included in this thesis

Drozdowska, B. A., Singh S., & Quinn, T. J. (2019). Thinking about the future: A review of prognostic scales used in acute stroke. *Frontiers in Neurology*, 10: 274.

Quinn, T. J., & Drozdowska, B. A. (2019). Stroke prediction and the future of prognosis research. *Nature Reviews Neurology*, 15(6), 311-312.

Elliott, E., Drozdowska, B. A., Taylor-Rowan, M., Shaw, R. C., Cuthbertson, G., & Quinn, T. J. (2019). Who is classified as untestable on brief cognitive screens in an acute stroke setting? *Diagnostics*, 9(3), 95.

Shaw, R., Drozdowska, B. A., Taylor-Rowan, M., Elliott, E., Cuthbertson, G., Stott, D., & Quinn, T. J. (2019). Delirium in an acute stroke setting, occurrence, and risk factors. *Stroke*, 50(11), 3265-3268.

Taylor-Rowan, M., Cuthbertson, G., Keir, R., Shaw, R., Drozdowska, B. A., Elliott, E., Stott, D., & Quinn, T. J. (2019). The prevalence of frailty amongst acute stroke patients, and evaluation of method of assessment. *Clinical Rehabilitation*, 33(10), 1688-1696.

Taylor-Rowan, M., Keir, R., Cuthbertson, G., Shaw, R., Drozdowska, B. A., Elliott, E., Evans, J., Stott, D., & Quinn, T. J. (2019). Pre-stroke frailty is independently associated with post-stroke cognition: A cross-sectional study. *Journal of the International Neuropsychological Society*, 25(5), 501-506.

Conference poster presentations

- European Stroke Conference (ESOC) 2019, Milan
- Alzheimer's Research UK Scotland Network Centre annual meeting 2019, St Andrews
- St Mungos Glasgow Royal Infirmary Research Symposium 2019, Glasgow
- UK Stroke Forum (UKSF) 2018, Telford
- British Geriatrics Society (BGS) Spring Meeting 2018, Nottingham

Abbreviations

4AT: 4 A's Test

A-Levels: Advanced Levels

AD8: Aging and Dementia-8

AIC: Akaike information criterion

AIREN: Association Internationale pour la Recherche et l'Enseignement en Neurosciences)

AMT-10: Abbreviated Mental Test

APOE: Apolipoprotein E gene

APPLE: Assessing Post-stroke Psychology Longitudinal Evaluation study

AUROC: Area under the receiver operating characteristic curve

BIC: Bayesian information criterion

BMI: Body mass index

BFI: Brief Fatigue Inventory

C-statistic: Concordance statistic

CAM-ICU: Confusion Assessment Method for the Intensive Care Unit

CDR: Clinical Dementia Rating Scale

CERAD: Consortium to Establish a Registry for Alzheimer's Disease

CESD-R: Centre for Epidemiologic Studies Depression Scale Revised

CFI: Comparative Fit Index

CHAID: Chi-square Automatic Interaction Detection

CHARMS: Checklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies

CI: Confidence interval

CNS: Canadian Neurological Scale

CSN: Canadian Stroke Network

COVID-19: Coronavirus Disease 2019

COWAT: Controlled Oral Word Association Task

CRF: Case report form

DISCs: The Depression Intensity Scale Circles

DSM: Diagnostic and Statistical Manual of Mental Disorders (-4 and -5 denote 4th and 5th editions, respectively)

EQ-5D: EuroQol – 5 Dimension

FIM: Functional Independence Measure

FIML: Full information maximum likelihood

FiND: "Frail Non-Disabled" Instrument

GAD-2: Generalised Anxiety Disorder Scale 2-item

GCS: Glasgow Coma Scale

GCSE: General Certificates of Secondary Education

GDS-15: Geriatric Depression Scale, 15 item

GMM: Growth mixture modelling

GRoLTS: Guidelines for Reporting on Latent Trajectory Studies

HER-2: Human epidermal growth factor receptor 2

IBM: International Business Machines Corporation

ICD: International Classification of Diseases

ICD-10NA: International Classification of Diseases, tenth revision: Neurological Adaptation

IQCODE: Questionnaire on Cognitive Decline in the Elderly

IQR: Interquartile range

KM: Kris McGill

LCGA: Latent class growth analysis

LDST: Letter-Digit Substitution Test

LISTEN: Loneliness Intervention using Story Theory to Enhance Nursing-sensitive Outcomes study

M: Mean

Mdn: Median

MMSE: Mini-Mental State Examination

MoCA: Montreal Cognitive Assessment

MOS-SSS: Medical Outcomes Study Social Support Survey

mRS: Modified Rankin Scale

NHS: National Health Service (UK)

NHS GGS: National Health Service Greater Glasgow & Clyde

NICE: National Institute for Health and Care Excellence

NIHSS: National Institutes of Health Stroke Scale

NINDS: National Institute for Neurological Disorders and Stroke

NPI-Q: Neuropsychiatric Inventory Questionnaire

OCS: Oxford Cognitive Screen

OCSP: Oxfordshire Community Stroke Project

OR: Odds ratio

PHQ-2: Patient Health Questionnaire – 2

PHQ-SADS: Patient Health Questionnaire – Somatic, Anxiety and Depressive Symptom Scales

PIS: Participant Information Sheet

PRECiS: Patient Reported Evaluation of Cognitive Status

PRISMA: Preferred Reporting for Systematic Review and Meta-Analyses

PROBAST: Prediction model Risk of Bias Assessment Tool

PROGRESS: Prognosis Research Strategy

RECREATE: Reducing Sedentary Behaviour after Stroke study

RMSEA: Root Mean Square Error of Approximation

SADQ-H 10: Stroke Aphasic Depression Questionnaire for patients in hospital, 10-item

SD: Standard deviation

SEM: Structural Equation Modelling

SF-SIS: Stroke Impact Scale Short Form

SOFA-Max: Sequential Organ Failure Assessment, maximum score

SPSS: Statistical Package for the Social Sciences

SRMR: Standardised Root Mean Square Residual

SSA-BIC: Sample size-adjusted Bayesian information criterion

STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

TIA: Transient ischaemic attack

TICS-M: Modified Telephone Interview for Cognitive Status

TLI: Tucker-Lewis Index

TQ: Terence Quinn

WHO: World Health Organisation

WLSMV: Weighted least squares mean and variance-adjusted estimator

WWI: World War I

Chapter 1 Introduction

In this chapter, I present concepts that are fundamental to my thesis, with a focus on cognition and stroke. After describing cognitive disorders in terms of their clinical features, and their prevalence and impact following stroke, I give special consideration to the topic of assessing cognitive function in healthcare and research settings. I regard this issue as a cornerstone of investigations into post-stroke cognition, including those within the area of prognosis research, which will be the theme of my next chapter.

1.1 Cognitive change

Cognition is central to how we perceive, understand, and interact with the world, and as such - central to our being. It encompasses many functions, from processing and interpretation of sensory stimuli, through remembering events and information, to use of language and complex operations on abstract concepts, such as in mathematical problem-solving or forming of philosophical doctrines. An obvious reflection is that it takes time for the full scope of cognitive abilities to become attainable for us, and so perhaps it is most intuitive to associate the term “cognitive change” with the dynamic development that occurs in childhood. Cognition, however, changes throughout the entire lifespan.

Some abilities, involving knowledge and skills acquired and consolidated through experience, education, and cultural influences (e.g. vocabulary, or familiarity with historical or geographical facts) may continue to improve into old age (1-3). Other abilities, relating to processing and learning of new information, and applying reasoning and problem-solving skills in a relatively unfamiliar context, begin to decline from the age of around 30 (1-3). This is recognised as an inherent part of “healthy aging”.

However, in some cases, either due to the type of functions that deteriorate, or the speed and extent of decline, change is indicative of a cognitive disorder, that is, a deficit in cognition that goes beyond what is attributable to the normal aging process. These problems are recognised as mild cognitive impairment or a

mild neurocognitive disorder (as referred to in the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders [DSM-5]) where they do not preclude an individual from living independently, although the performance of some activities is slower and more effortful than it was previously, entails more errors, and requires use of compensatory strategies (4-7). Conversely, in its severe form, cognitive impairment compels the individual to rely on the help of others to manage the requirements of daily life. This is a key aspect of dementia or, as termed more broadly, a major neurocognitive disorder (6, 8).

With influences present at every stage of life, multiple variables are associated with cognitive decline and acquired cognitive impairment, including: genes, socioeconomic background, education, occupation, environmental exposures, social engagement, lifestyle choices, and health-related conditions (9-11). Among the latter, stroke – of which there are over 80 million prevalent cases worldwide (12) – has been consistently demonstrated as a major contributing factor, at least doubling the risk of developing dementia (13).

1.2 Stroke

The World Health Organisation (WHO) and National Institute for Health and Care Excellence (NICE) traditionally define stroke as a clinical syndrome of vascular origin, characterised by sudden onset of rapidly developing signs of a focal or global disturbance of cerebral function, lasting at least 24 hours or leading to death (14, 15). Episodes where neurological dysfunction is suspected to be caused by ischaemia, yet symptoms resolve within 24 hours and there is no evidence of acute infarction, are identified as a transient ischaemic attack (TIA) (15).

In recent years, however, these classic definitions of stroke and TIA have been increasingly recognised as outdated and of limited usefulness in a clinical setting (16, 17). A statement from the American Heart Association and American Stroke Association highlighted the importance of accounting for advances in science and technology for defining stroke and its diagnosis (16). Neuroimaging was discussed as a valuable source of objective evidence of central nervous system infarction, needed to supplement findings from clinical observation. Specifically, use of

neuroimaging can help to: i) characterise a lesion (its location, size, shape, and extent); ii) differentiate between an ischaemic (resulting from the obstruction of an artery) and a haemorrhagic (a focal collection of blood from the rupture of a vessel) stroke; iii) identify a silent brain infarct, where there is no history of acute neurological dysfunction attributable to a lesion; iv) differentiate between a stroke and a stroke mimic (e.g. brain tumour, migraine); and v) differentiate between a stroke and a TIA.

The role of neuroimaging in fulfilling the latter objective has become particularly emphasised through increasing controversy over applying a time-based criterion alone to differentiate between stroke and TIA. As studies have indicated that permanent infarction occurs in around one third of patients with symptoms lasting under 24 hours, it is argued that failing to recognise the limitations of this traditional rule of thumb may have led to misdiagnoses and thus inappropriate treatment (16, 18). In the context of cognition, it is moreover relevant that there is some evidence to suggest an association between TIA and longer-term cognitive problems (19, 20). With this in mind, while I focus on stroke, many of the concepts I discuss in this chapter are also relevant to cases of TIA.

1.3 Cognitive disorders following stroke

1.3.1 Syndromes, prevalence, and impact

The neurological damage caused by stroke, as well as the medical complications that may follow, entail a risk that is rarely considered in relation to the general home-dwelling population – that of delirium. One in four stroke survivors are likely to be affected by this condition during the first days post-ictus. Delirium is characterised by disturbed attention and cognitive function, with a sudden onset and fluctuating course (21). Although generally considered a transient state, the effects of delirium are not always reversible, with potential severe implications in terms of length of hospital stay, disability, subsequent cognitive decline, and mortality (22). Delirium can be particularly difficult to recognise following stroke, due to the likely co-occurrence of pre-existing and/or acute cognitive impairment.

Estimates suggest that around 10% of individuals have dementia prior to index stroke, while within just the six months that follow, another 10% are likely to develop new dementia (23). At least three times as many stroke survivors are indicated to have mild global cognitive impairment, deficits in single cognitive domains, or subjective cognitive complaints (24-27). Indeed, a number of studies suggest that the majority of the stroke population will be affected by some form of short-term or persisting cognitive problems (25, 26), with potentially severe, adverse implications for individual outcomes. These include reduced functional gains from rehabilitation, inability to return to work, limited social participation, dependency in activities of daily living, increased risk of mood disorders, and poor quality of life (28-32).

In addition to personal losses, post-stroke cognitive difficulties may affect the well-being of family members and increase caregiver burden (33-35). There are further ramifications at a societal level, with the presence of cognitive disorders associated with increased healthcare costs, stemming from longer initial hospitalisation, greater risk of later readmission and institutionalisation, and increased use of outpatient and home-based support (36).

1.3.2 Improving cognitive function following stroke

Given this high prevalence and extensive, detrimental impact of cognitive disorders, it seems unsurprising that in relation to life after stroke, finding ways to improve cognitive function has been determined as a number one research priority, in consensus by stroke survivors, caregivers, and health professionals (37). To date, there is limited evidence to support current interventions (38). One of the main postulated approaches involves use of strategies found to contribute to favourable stroke outcomes in general, including: treatments in the stroke unit to prevent acute complications, early rehabilitation, and pharmacological and lifestyle interventions for reducing the burden of cardiovascular risk factors and preventing recurrent stroke (39).

Importantly, however, findings regarding the impact of such strategies on improving cognitive outcomes, specifically, seem thus far inconclusive. For example, while some studies have reported favourable effects for thrombolytic therapy, active blood pressure lowering, and lipid lowering, others indicated

neutral results (40-46). Moreover, two randomised controlled trials demonstrated no benefit of multicomponent interventions to reduce cardiovascular risk burden for the improvement of post-stroke cognition; one involved pharmacological strategies and lifestyle modifications, the other – only non-pharmacological strategies, although this included encouraging compliance with prescribed medication (47, 48).

Similarly, there are uncertainties regarding the effectiveness of interventions that directly target post-stroke cognitive function through use of restorative and compensatory strategies, i.e. cognitive rehabilitation. The authors of a recent systematic review and meta-analysis made an encouraging finding that such approaches had a small (across 7 controlled studies) to moderate (across 13 pre-post studies) positive effect on post-stroke cognition (49). At the same time, however, all included studies were rated as being of low quality and high risk of bias. Further, there was no evidence to indicate whether the observed effects were long-lasting, with only three studies including a follow-up assessment, the latest conducted one month after completing the tested strategy.

In view of the above, there is a need to continue developing and/or tailoring interventions to improve cognitive function following stroke. Such endeavours involve a multi-stage process, from understanding the distinct features and natural history of post-stroke cognitive problems, through identifying their determinants and opportunities for modifying their course or manifestation, to testing and eventually implementing person-tailored interventions in routine clinical practice. There is one component that is essential to success at any stage – the ability to accurately detect a post-stroke cognitive disorder.

Assessing cognitive function following stroke is in itself a challenging task. Before I refer to issues around the selection, feasibility, and applicability of assessment approaches, it seems important to consider a more fundamental challenge – defining vascular cognitive syndromes for the purpose of diagnosis.

1.3.3 Challenges to diagnosing neurocognitive disorders following stroke

Traditionally, endorsed definitions of dementia primarily encapsulated features of Alzheimer's disease, treating impairment in memory and continuing decline in cognitive function as core criteria for diagnosis (50). However, cognitive disorders of vascular origin are often characterised by deficits in attention, speed of information processing, and executive function, and do not necessarily progress over time (51, 52).The recognition of this gap in the conceptual approach to diagnosis has prompted much debate over developing and implementing definitions of "vascular cognitive impairment" and "vascular dementia" for both clinical and research use, with controversies regarding diagnosis, classification, and terminology present to this day (50, 53, 54).

Overall, proposed criteria focus on the presence of a relationship between cognitive deficits and cerebrovascular disease, evidenced by focal neurological signs on examination (e.g. hemiparesis, dysarthria) and/or neuroimaging findings (presence of infarcts, lacunes, and white matter lesions) (55-57). A degree of uncertainty is inherent – pure vascular dementia is considered rare in an older adult population, with brain lesions of vascular origin possibly contributing to or merely co-occurring with the effects of ongoing neurodegenerative processes (38, 58). While there have been attempts to validate suggested diagnostic criteria for vascular cognitive disorders (59-61), consensus recommendations are still lacking, and a definite diagnosis can only be reached through including findings from post-mortem investigations (58).

In relation to stroke, these challenges in the application of diagnostic labels have led some experts in the field to adopt a more pragmatic approach to defining subsequent cognitive disorders. Specifically, the term "post-stroke dementia" has been proposed for any dementia that develops following stroke, without imposing specific criteria regarding the underlying neuropathological process(es) (38). Within this framework, post-stroke dementia constitutes a sub-type of: vascular cognitive impairment, vascular dementia, and post-stroke cognitive impairment. In view of the practical advantages, when referring to post-stroke cognitive disorders throughout my thesis, I apply a similar approach –

assuming a temporal relation between stroke and cognitive change, while recognising that the former does not necessarily act as the predominant causal factor.

1.3.4 Assessment of post-stroke cognitive function

1.3.4.1 Identifying cognitive disorders in a clinical context

Early screening for cognitive deficits is recommended in all stroke survivors, although currently there is no consensus regarding an optimal approach to assessment (62). While the specific content of screening measures varies, they are generally designed to be brief and require relatively little training for correct administration (63). Typically, at least a few different cognitive functions are assessed, such as language, attention, or learning memory. However, rather than performance on individual tasks, it is the sum score that is of particular focus, serving as an indicator of global cognitive status. Deciding on the most appropriate screening tool from the many that are currently available will, at least in part, depend on the circumstances and setting.

In the first hours and days following stroke, attending to a patient's medical needs is likely to be prioritised over an assessment of cognitive function (62). Nonetheless, in interest of the former, it is necessary to recognise delirium to initiate appropriate interventions as early as possible (64, 65). Validated measures such as the 4 A's Test (4AT) (66, 67) and the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) (68) combine relatively good accuracy with high feasibility in the stroke context, taking under five minutes to complete, and being suitable for patients with motor, visual, and speech impairments (69).

In addition, some insight into a patient's cognitive state in the hyperacute phase can be gained through initial neurological examinations. Despite not representing cognitive screening tools per se, the Glasgow Coma Scale (GCS) (70), National Institutes of Health Stroke Scale (NIHSS) (71, 72) or Canadian Neurological Scale (CNS) (73) all include items relevant to cognitive function (e.g. orientation, speech comprehension and production).

Once a patient is medically stable, a more comprehensive assessment can be considered, although feasibility is likely to remain a key concern (62). On one side, there is the issue of limited resources in terms of time, space, staffing and funds. On the other, there are factors relevant to the patient's condition, such as becoming easily fatigued, experiencing distress, or being unable to complete certain cognitive tasks due to acquired deficits (e.g. drawing a clock with upper limb weakness, or object naming with a severe visual deficit) (74).

The latter is an important issue even in relation to some of the most widely used screening tools for identifying cognitive impairment (75), such as the Mini-Mental State Examination (MMSE) (76) or the Montreal Cognitive Assessment (MoCA) (77). As these measures were not specifically designed for use in a stroke setting, it is important to interpret findings with caution, and potentially consider adjusting the originally determined cut-off values, so that chosen tools are more "fit for purpose" (78, 79). However, deciding on optimal cut-offs is in itself a considerable challenge, due to a common trade-off between two key test properties – sensitivity and specificity (80).

Sensitivity refers to the proportion of individuals who have a certain condition that are correctly identified as having it (81, 82). Specificity relates to the proportion of individuals who while not having the condition are correctly classified as not having it (81, 82). Regardless of which test property is favoured, some negative consequences are to be anticipated. Poorer sensitivity entails a higher probability of genuine cognitive disorders being missed, and thus some individuals not receiving the follow-up and support they need; poorer specificity may lead to more individuals experiencing unnecessary distress and stigma due to an erroneous indication of an impairment, as well as misallocation of healthcare resources.

Given the limitations of cognitive screening tools, the implications of misidentifying the presence vs. absence of cognitive disorders, and the dynamic process of post-stroke recovery, in a clinical context, findings from early, brief testing are not recognised as definitive or fulfilling a diagnostic purpose (83). Rather, they may serve to monitor for potential change in function, and inform initial care plan decisions, particularly regarding whether a more detailed cognitive assessment, involving use of a comprehensive neuropsychological test

battery, is recommended (83, 84). Such assessments are considerably longer, and require specialist training to conduct and interpret, entailing increased test burden for stroke survivors and health system costs. For these reasons, comprehensive neuropsychological assessments are offered to only selected individuals, with suspected cognitive problems (62, 75).

Neuropsychological test batteries, comprising of multiple tasks, are designed to determine the presence and severity of deficits in specific cognitive functions, relative to population normative data (83, 85). This provides a more in-depth understanding of an individual's unique cognitive profile, which may involve either single domain (e.g. within executive function) or multidomain (e.g. within learning memory and executive function) impairment. Findings from such an assessment can serve an important role in tailoring rehabilitation interventions according to individual needs (83).

Although comprehensive neurological assessments are recognised as the gold standard for identifying and characterising cognitive deficits following stroke (62, 75), they are not the only source of information that needs to be considered for a clinical diagnosis of a mild or major neurocognitive disorder. As indicated in the DSM-5 criteria (6), other key aspects include a subjective concern of the individual or informant over a decline in cognitive function from previous status, and the effect of cognitive deficits on performing daily activities (8). It is important to note that these criteria are not stroke-specific, which can entail certain challenges. Specifically, it may be difficult to discern the impact of cognitive problems on day-to-day functioning and independence, as following stroke they are often accompanied by physical impairments, constituting another plausible cause of disability (38).

1.3.4.2 Identifying cognitive disorders in research

Reflecting clinical practice guidelines, current research recommendations advocate for all stroke trials to include an assessment of cognitive outcomes (86). This is suggested to involve a two-step procedure: i) an informant-based assessment to ensure study groups are well-matched in terms of pre-stroke cognitive status, and ii) a neuropsychological assessment, conducted between three to six months post-stroke. Regarding the latter, it has been proposed that

at a minimum the assessment should comprise a screening measure for global cognitive impairment (such as the MoCA) and additional brief tasks targeting attention and executive function (86).

However, across the whole landscape of research around post-stroke cognition, there is much variability in assessment methods, suggesting that the choice of approach is likely guided by individual study aims, the setting, and potential challenges and limitations (87). Providing a feasible solution, a screening tool will often be the only measure used to discriminate between individuals with and without a post-stroke cognitive disorder (87, 88). Similarly, to at least partially reduce the resource burden associated with large multicentre follow-up studies, investigators may opt to conduct remote assessments (e.g. over the telephone or online), or even retrospectively derive data regarding presence of cognitive disorders from medical records. Further, in some studies the use of objective assessment methods will be replaced by self-report or informant-based questionnaires to determine cognitive outcomes (26).

While in many cases a given assessment approach is simply viewed as a means to an end (the latter being the detection of a post-stroke cognitive disorder), in some research it is the method itself that is of primary interest. An example is provided by studies that investigate the accuracy of cognitive tests developed for use in the general population when applied specifically to stroke (69, 79). Other studies may focus on the development and performance of novel and/or stroke-specific methods. In recent years, this has not only involved investigations into direct assessments of cognitive function (e.g. 89, 90), but also the use of biomarkers, including metabolic, genetic, inflammatory and neuropathological factors (38, 91-93).

Such studies, determining how well a measure can discriminate between individuals with and without a certain condition as compared to a reference standard, represent diagnostic research (94). This is one of two areas falling under the broader scope of prediction research, at the centre of which is estimating the probability of something presently unknown (95). The second area – prognosis research – is the focus of my next chapter, as a topic that is of key relevance to this thesis.

1.4 Summary

Stroke significantly increases the risk of cognitive problems, which may have a profound impact on affected individuals, their families, and health and social care structures. To date, there is limited evidence in support of specific interventions to improve post-stroke cognitive function in the longer-term. Accurate detection of cognitive problems is an essential component of investigations that could eventually lead to the development and implementation of effective therapeutic strategies.

There are, however, many challenges to assessing cognitive function following stroke, including limited resources and a consequent need for their prioritisation, difficulties experienced by stroke survivors in participating in assessments due to acquired impairments, and limited applicability of endorsed cognitive measures developed for use in the general population. Diagnostic studies contribute to the development of more appropriate and accurate methods for identifying post-stroke cognitive disorders, including tests that accommodate common deficits (e.g. visual impairments, hemi-spatial neglect, or aphasia), as well as the use of novel biomarkers.

Chapter 2 Prognosis research: objectives, components, and advances in the area of post-stroke cognition

Having described concepts that are fundamental to my thesis as relating to cognitive function and stroke, I proceed to describing the context within which I address issues of interest – that of prognosis research. In this chapter, I firstly provide an overview of the types of investigations that prognosis research encompasses, with emphasis on their specific aims and relevant methodological approaches. I then present findings from prognosis studies focusing on post-stroke cognitive outcomes, concluding with a reflection on current advances in this area of research.

2.1 What is prognosis research?

While diagnostic studies focus on conditions that are already present, but not yet detected, prognosis studies focus on the development of future outcomes (95). This sets a unique purpose for prognosis research – to address this risk of an unfavourable outcome so that future health can be improved (96). As outlined within the PROGnosis RESearch Strategy (PROGRESS) (96), the types of investigations that contribute to achieving this goal can be divided into four themes: i) fundamental prognosis research, ii) prognostic factor research (also referred to as predictor finding research), iii) prognostic model research, and iv) stratified medicine research. Evidence collected under earlier themes serves to inform investigations under subsequent ones.

2.2 Fundamental prognosis research

This first theme in prognosis research describes the natural history and clinical course of a health condition (97). Its aim is to estimate the "baseline risk" for a particular outcome in a specific population, which typically involves an observational approach (98). Studies within this theme allow to answer such questions as: "On average in the UK, what is the risk of death within five years of heart failure?" or "How likely is the development of delirium among acute stroke unit inpatients?". Fundamental prognosis research also provides grounds

for comparing baseline risks across different clinical contexts (96), for example, to determine whether five-year mortality rates following heart failure differ between hospitals, countries or across decades. Such findings may provide important insights into how specific health policies and aspects of local routine care shape patient outcomes, and aid in identifying potential targets for system improvement.

2.3 Prognostic factor research

Clinical contexts constitute only one of many types of variables that can be associated with the development of a future health-related outcome. Identifying these variables is the focus of the second theme – prognostic factor research (99). Different approaches can be applied to investigate this topic, with the chosen method likely to impact on how much confidence can be placed in the findings obtained. Evidence from studies estimating the correlation between a single presumed prognostic factor (candidate predictor) and an outcome of interest, with both assessed at the same point in time (cross-sectional design), is considered suitable only for hypothesis-generating purposes (97). This is due to the significant limitations such an approach entails.

Firstly, the nature of the association between two variables can change over time. For example, evidence suggests that in late life high blood pressure is associated with a reduced risk of dementia (100). Recognising this relationship would not, however, be applicable to estimating risk of future dementia in mid-life, as here the association is reversed – high blood pressure increases the risk of poor future cognitive outcome (100).

Secondly, in a univariable analysis, the effects of other variables are unaccounted for, which can lead to spurious findings, and thus erroneous conclusions regarding studied associations. This issue is illustrated by a commonly observed relationship between female sex and unfavourable health outcomes. In some cases, the estimated association reflects a genuine phenomenon, driven by biological or socio-cultural mechanisms, such as the finding that obese women are at greater risk of heart failure than obese men (101). However, such associations may also be driven by age – as women on

average live longer than men (102), the oldest participants in a study sample are more likely to be female.

For the above reasons, research evidence regarding prognostic factors is considered more reliable where it stems from multivariable analyses of longitudinal data (103). This requires conducting observations on at least two different timepoints, where information regarding the outcome is preferably collected later than data on multiple candidate predictors. The latter are simultaneously included in an analysis, allowing to quantify their independent associations with the outcome, that is, the direction and strength of their relationship when the influence of other, potentially relevant factors is controlled for.

It is important to note that while similarities exist, such an investigation is not equivalent to aetiological research. The latter aims to explain the cause of an outcome, while the purpose of prognosis research is solely to predict it; as such, there is no need to determine whether observed associations are causal or non-causal (103, 104). Nonetheless, identifying causal factors is also recognised as being of particular value in the context of prognosis research. One reason is that causal factors may serve as targets for intervention, assuming that it is possible to improve the outcome through their modification. Where the variable is non-modifiable, e.g. as generally is the case for genetic factors, it may still be highly relevant to predicting differential treatment responses – a property of key interest in the area of stratified medicine (105).

Moreover, where the relationship between a factor and outcome is causal, based on biological (or other) pathways, it is more likely to be consistently present across different populations, entailing enhanced generalisability of research findings (106). While the latter is of concern in any study, specifically for prognostic model research, generalisability is a quality that essentially determines its value (107).

2.4 Prognostic model research

Prognostic model research involves three phases: model development, external validation with potential updating, and investigation of the impact of implementing a prognostic model in clinical practice (106). Conceptually, this multi-stage process is not intended to be finite – it is argued that model validation and updating should be ongoing, ensuring applicability throughout varying contexts and changing times. This notion, combined with a need for relatively large sample sizes from different settings, and longitudinal study designs, entails that prognostic model research is time-consuming and resource-intensive. As such, it seems unsurprising that, in many cases, research work relating to a prognostic model does not extend past its initial development (106).

2.4.1 Model development

In prognostic model development studies, variables identified as associated with an outcome through prognostic factor research are combined to estimate the probability for an individual to develop that outcome (103, 106). It is noteworthy that a prognostic model may be identical to a multivariable model used for the exploration of candidate predictors. Moreover, both types of investigations – relevant to prognostic factors and prognostic models – are often conducted in a single study. Consequently, it may be difficult to distinguish between the two prognosis research themes.

A key difference lies in the focus of a study (97). In prognostic factor research, of particular interest is quantifying the relationship between individual variables and a future outcome. In prognostic model research, the focus is on identifying a set of factors, which collectively can accurately estimate the likelihood of the future outcome. The final result of such work can be referred to as a clinical or risk prediction model, or a prognostic rule, index or score (106). It may be presented directly in the form of a multiple regression equation:

$$Ln[p/(1-p)] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \ldots$$

This formula applies to logistic regression, used for binary outcomes, where Ln is a natural log, p – the estimated probability of the future outcome, α – the intercept, β_N – the estimated coefficient that reflects the association between a predictor and the outcome, and X_N – the value of the predictor.

Alternatively, based on the specified equation, researchers may aim to develop a simplified, more user-friendly tool. This will often involve rounding up estimated predictor coefficients to integers to produce a risk score, where points are assigned according to the presence/value of included risk factors and then summed. The total score can then be related to pre-determined cut-offs, indicating the associated level of risk for an unfavourable outcome.

An example of such an approach is illustrated by ASTRAL – a risk score developed for the prediction of an unfavourable functional outcome (functional dependency) following acute stroke (108). The incorporated predictors and scoring system are presented in Table 2-1. As reported by the authors, example scores of 23, 31, and 38 correspond to a 20%, 50% and 80% likelihood of an unfavourable outcome, respectively.

Table 2-1 ASTRAL variables and scoring system.

Variable	Level/category	Score
Acute glucose	<3.7 or >7.3 mmol/L	1
Age	Per every 5 years	1
Any stroke-related visual field defect	Yes	2
Level of consciousness	Decreased	3
Symptom onset to treatment time	>3 hours	2
Stroke severity as per NIHSS	Per every point	1

The same publication also demonstrated another approach to framing a prognostic rule – through use of a graphical representation. Here, this involved a display of multiple, coloured charts. Different charts applied to different combinations of risk factors, while each colour corresponded to a different probability of the future outcome (e.g. light blue: 30 to 39%, dark red – 80 to 89%).

2.4.1.1 Assessment of prognostic model performance

The process of prognostic model development should also involve an assessment of the model's performance (106). Although some indication is provided by the amount of variance in the outcome explained by the chosen set of predictors, it is also essential to estimate model discrimination and calibration (109). Discrimination relates to the ability of a prognostic model to distinguish between individuals who develop the outcome and those who do not (110). The primary method for assessing this property is the concordance statistic (C-statistic), which for binary outcomes corresponds to the area under the receiver operating characteristic curve (AUROC) (109, 111). Possible values range from 0.50, indicating a discriminatory ability equivalent to chance, to 1.00, indicating perfect discrimination (112).

Calibration refers to the level of agreement between observed and predicted outcome probabilities (113). It is preferably assessed through inspection of calibration plots (113). The results of the Hosmer-Lemeshow test (114) may also be reported to complement graphical evaluation, yet as a stand-alone measure it is considered inappropriate (despite often being used in this capacity) (111). The test has been recognised to have limited power to detect poor calibration, is oversensitive in large samples, and cannot inform about the direction of miscalibration (107)

The assessment of discrimination and calibration is a key component of model validation, which at the stage of development is either apparent or internal (106, 115). In apparent validation, performance is evaluated directly in the dataset that was used for model development. Findings suggesting low prognostic ability immediately indicate potential issues in the derivation process. At the same time, encouraging results are considered as insufficient evidence of a model's prognostic value. This is because obtained performance estimates are likely to be overoptimistic, as the coefficients reflecting predictor-outcome associations were optimised for that specific data (109, 116).

Methods applied for internal validation are intended to at least partially correct for this issue. One technique – split-sample – involves dividing the initial study sample into separate development and validation cohorts (117). This approach,

however, is criticized for the loss in statistical power it inherently entails. Therefore, current recommendations advise employing data re-use techniques, one of which is an extension of the split-sample method – cross-validation (109). In this case, the sample is randomly divided into a number of equal-sized groups (the number being specified by the researcher) and one group serves as a validation set, while all remaining groups are involved in model development (117). The process is repeated multiple times until every group served as a validation set once. Model performance is estimated as an average across all repetitions. For example, a 10-fold cross validation involves 10 repetitions, where for each one a different group, constituting 10% of the overall sample, will be used for validation, and the remaining 9 groups, constituting 90% of the sample, will be used for model development.

A second data re-use technique, currently recognised as the most efficient approach to internal validation, is bootstrap resampling (117). Bootstrapping mimics the process of generating a sample from an underlying population. Random samples are drawn with replacement from the initial study cohort, generating bootstrap samples of the same size. Some study participants may not be included in a specific bootstrap sample at all, while others may reappear in a generated dataset multiple times. In Table 2-2, I presented a basic example of what bootstrap samples may look like, using a single variable (height) for ten participants.

Table 2-2 Illustration of four bootstrap samples drawn with replacement from the height of ten participants.

Sample (N = 10)	Participant height (cm)
Original	162, 183, 177, 172, 155, 164, 171, 161, 188, 158
Bootstrap 1	183, 155, 164, 171, 188, 158, 155, 183, 158, 171
Bootstrap 2	162, 177, 172, 155, 171, 177, 177, 162, 155, 172
Bootstrap 3	172, 164, 161, 188, 158, 162, 155, 183, 161, 164
Bootstrap 4	183, 155, 155, 183, 171, 171, 188, 155, 172, 183

The prognostic model is developed in the bootstrap samples and then validated in the original dataset (118). The decrease in model performance between the bootstrap and the original samples provides an estimate of optimism –

subtracting it from the apparent performance indicates what predictive ability can be expected from the model in future cohorts. However, it is only through external validation that the latter can be truly determined.

2.4.2 External validation

Before describing this stage in prognostic model research, it is important to note that some researchers distinguish a form of performance evaluation that in terms of stringency is recognised as intermediate between internal and external validation (119, 120). This is temporal validation, where the model is tested in a new sample of participants, recruited at a later period from the same setting (e.g. hospital site) as the development sample. Despite including different individuals, the case-mix is plausibly similar to the original cohort, as the validation sample is still drawn from the same underlying population. As follows, there will remain doubts as to how well the model will perform in a different context.

Quantifying the latter is the goal of external validation, which allows to verify the generalisability of a prognostic model (120, 121). The extent of any conclusions will depend on the form of external validation. A common approach, referred to as geographical validation, essentially involves applying the model to a sample from a different location (122). This in itself encompasses a range of options, from testing at a different centre within the same region, to applying the model in a sample from a different country (or continent), entailing effects of cultural, ethnic and health system dissimilarities.

Another approach, termed methodologic validation, allows to determine whether prognostic accuracy is maintained despite using different methods of collecting data and inconsistencies in operationalisation of variables (121). Any challenges in this regard are particularly likely to be exposed when the model is tested by different, independent investigators, with some factors plausibly more robust to variability in assessment than others. For example, while hardly any inconsistencies can be expected in the measurement of age, concluding on whether a participant has hypertension may differ depending on whether one relies on a diagnosis present in existing medical records, prescription of antihypertensive medication, or acute measures of blood pressure.

External validation can further focus on differences regarding population characteristics (121). These can pertain to the prevalence of risk factors, as well as attributes of the condition that constitutes the outcome – its incidence, degrees of severity across the sample, or its clinical course. Finally, the generalisability of a prognostic model can be evaluated in relation to differences in follow-up duration, that is, the length of the interval between a specific baseline event (e.g. having a stroke or receiving a cancer diagnosis) and outcome assessment (121).

It is important to note that while these distinctions provide a useful framework for describing and understanding the multiple aspects of prognostic model generalisability, it is likely that an external validation will differ from the model development study on more than one account. For example, if a prognostic model is developed in a sample of hospital inpatients admitted with myocardial infarction in Scotland, subsequent validation in a similar population in China (i.e. with the same diagnosis, recruited from a comparable clinical setting) would likely entail more than just geographical diversity, but also variability in prevalence of risk factors, their measurement, and the clinical course of the condition (e.g. due to differences in routine clinical care).

2.4.3 Prognostic model updating

If on external validation the performance of a prognostic model is found to be unsatisfactory, it may be potentially improved through use of an updating method. This can simply involve model recalibration. In cases where a decline in prognostic accuracy is attributed to a difference in the incidence of an outcome as compared to the development sample, the average predicted probability can be adjusted to align with the currently observed event rate (equivalent to updating the intercept) (123, 124). Where it appears that estimated predictor coefficients in the original model were overfitted, these can also be adjusted – by a single adjustment factor, assuming that the relative effects of predictors are similar, but the absolute effect sizes ought to be either larger or smaller (equivalent to updating the intercept and slope) (123, 124). To clarify, overfitting is a common issue, where maximising adherence of a developed model to the unique characteristics of a used dataset leads to increased model

complexity, overoptimistic initial performance estimates, and limited applicability to new samples (125).

If recalibration alone is insufficient to achieve an expected level of prognostic performance in a new sample, investigators may consider replacing or adding model predictors. The decision-making process can be guided by different rationales. One involves promoting methodologic generalisability, by substituting variables associated with high inter-rater variability with ones that can be measured more reliably (106). Another is based on a previously mentioned argument, favouring the inclusion of prognostic factors with a causal relationship with the outcome, and thus that are more likely to show consistent associations across different contexts.

Thirdly, the development and increasing availability of new techniques may lead to identifying novel prognostic factors for addition to existing models, or can provide a more accurate measurement of variables that had been previously considered, for example, as in the case of rating white matter changes based on magnetic resonance brain imaging as compared to computed tomography (126). It is also this rationale that seems to strongly justify the abovementioned recommendation for prognostic models to be continuously updated.

2.4.4 Prognostic model impact

Studies in this area aim to determine whether the implementation of a prognostic model leads to better outcomes than are achieved through routine clinical practice (127). Such investigations focus on improvements in individual outcomes or the cost-effectiveness of care, where change is attributable to the influence of prognostic model use on clinical decision-making (106). This is preferably assessed in randomised controlled trials (128).

A fundamental understanding in prognostic model impact research is that even for a model with excellent prognostic ability to be of benefit, it must first lead to a change in clinician behaviour (127). One important issue in this context relates to feasibility. Collecting predictor information may be recognised as too costly or time consuming in a routine care setting, or unacceptable to the target population (e.g. due to the invasiveness of a procedure), while the whole

process of estimating individual outcome risk may be viewed as too complex or confusing. Other concerns may be over practicing “cookbook medicine”, with prognostic rules considered too generic, and their use undermining a comprehensive, individual approach to patient needs (128, 129). Clinicians may moreover doubt the tools validity and accuracy, and assume that clinical judgement is overall superior (128).

2.5 Stratified medicine research

2.5.1 What is stratified medicine?

From both research and clinical perspectives, people tend to be grouped on the basis of sharing a few particular traits that are of importance in a specific context, e.g. women who are pregnant, children with asthma, or older adults with arthritis. However, in relation to many other characteristics, including individual risk factor profiles, there may be much heterogeneity within a conceptualised group. This can entail variability in the clinical course of a condition, the likelihood of particular future outcomes, and individual treatment response. Taking this into consideration is the cornerstone of stratified medicine, which aims to maximise the beneficial impact of healthcare through targeting interventions according to the clinical characteristics of specific patient subgroups (105).

At a basic level, stratifying the use of treatments may be guided by an individual’s absolute risk for an unfavourable outcome (105). When the relative effect of an intervention is found to be similar across all patients, the absolute reduction in the probability for the unfavourable outcome will be greatest for those who were initially at highest risk. For example, if a treatment is associated with a relative risk reduction in stroke incidence by one third (33%), for a person whose baseline risk of having a stroke was estimated at 60%, this will translate to an absolute risk reduction of 20%, while for a person with an initial risk at 15% - to an absolute risk reduction of around 5% (130). In cases where an intervention is costly or entails nonnegligible side effects, it may be offered only to those who are likely to benefit the most.

A second approach involves stratifying treatment based on the presence of a factor (or factors) that may influence the effect it has (105). This can relate to both potential benefits of an intervention, as well any harms. For example, there is strong evidence supporting use of oral anticoagulation in patients with atrial fibrillation to prevent embolic stroke, however, use of anticoagulants also predisposes to bleeding (131). While for many individuals the protective benefits will outweigh the risks, for those with a high baseline probability of bleeding, such as in cases of high alcohol consumption or liver disease, the danger of major haemorrhage may tip the scales in the opposite direction (132).

2.5.2 The role of prognosis research in stratified medicine

Prognosis research contributes to stratified medicine on multiple levels (105). Firstly, it may inform prioritisation of topics for investigation, for example, by indicating high heterogeneity in the clinical course and prognosis for a certain condition (as in systemic lupus erythematosus (133)), or significant interindividual differences in the metabolism of a particular drug (134). Secondly, prognosis research is relevant to both abovementioned approaches to treatment stratification: on one hand, building the necessary evidence base to accurately estimate the baseline risk of an unfavourable outcome, and on the other, leading to identification of prognostic factors that can predict differential treatment response.

Regarding the latter, in some cases decisions regarding the appropriateness of a specific treatment may be guided by the presence of one particular factor. An example of such a scenario is often illustrated by the discovery that trastuzumab significantly improves disease-free and overall survival in women with breast cancer who have a positive human epidermal growth factor receptor 2 (HER-2) gene status (135, 136). At the same time, compared to women with a negative HER-2 status, they benefit less from other, standard cancer treatments, and have a poorer baseline prognosis (without use of trastuzumab).

In other cases, therapeutic decision-making may be aided by use of validated prognostic rules, involving an assessment of multiple factors. Returning to the above example on atrial fibrillation management, as a first step, current recommendations suggest that patients are assessed for their risk of stroke

based on the CHA_2DS_2VASc score (131, 137, 138). Where the risk is very low, anticoagulation treatment may be withheld to avoid unnecessary exposure to adverse effects of medication, as well as incurring healthcare expenses. Where the risk is considerable, the next step is to assess the probability of bleeding using the HAS-BLED score (139). Patients at high risk of both stroke and bleeding are candidates for an alternative, invasive intervention, involving closure of the percutaneous left atrial appendage (140).

Finally, prognostic research may assess the impact of newly introduced approaches to treatment stratification. Similarly as in the context of prognostic model impact studies, such investigations can focus on: clinician adherence and changes to behaviour, including barriers and facilitators to implementing a specific approach; the influence on patient outcomes, including both beneficial and adverse effects; and cost-effectiveness.

2.6 Prognosis research into post-stroke cognitive outcomes

Although this may not have been evident from my formulation of the topic, I included some examples of research in this area in Chapter 1. The studies I referenced can be classed as pertaining to the theme of fundamental prognosis research, leading to such observations as: “one in four stroke survivors are at risk of developing delirium in an acute setting” (21) or “one in ten individuals are likely to develop new dementia within one year of stroke onset” (23).

The second statement, relating to post-stroke dementia, is derived from a highly influential publication (cited nearly 1500 times) that is also of particular relevance to the second prognosis research theme – identification of prognostic factors. This is a systematic review and meta-analysis from 2009 by Pendlebury and Rothwell, who reviewed 73 papers, involving a total of 7511 stroke survivors (23).

2.6.1 Predictors of post-stroke cognitive impairment

Alongside determining the incidence of both pre- and post-stroke dementia, the authors of the review quantified pooled effects for multiple risk factors, both pre-dating the stroke, as well as specific to it. Among relevant demographics

were older age, female sex, low educational attainment, and being of either black or Hispanic ethnic origin. In relation to prior health-related conditions, the authors determined associations with diabetes mellitus, atrial fibrillation, previous stroke, disability, prior cognitive decline, and neurological changes identified through imaging – leukoaraiosis (white matter abnormalities) and cerebral atrophy.

Regarding features of the index stroke, the risk of developing dementia was reported to increase with greater severity, greater infarct volume, left hemisphere lesions, and haemorrhages (as compared to ischaemic strokes). The review further indicated the relevance of acute symptoms and complications, including: aphasia, incontinence, seizures, confusion, hypotension, and hypoxic ischaemic episodes. Finally, individuals who either had multiple infarcts or a recurrent stroke were approximately 2.5 more likely to develop dementia.

In comparison, two recent reviews of prognostic factors for post-stroke cognitive impairment presented a narrower focus, mainly relating to novel biomarkers (141, 142). Reflecting a narrative approach, one publication specifically examined the evidence for inflammatory (e.g. C-reactive protein, interleukin 6 and 10), metabolic (e.g. homocysteine, retinoic acid), growth factor (e.g. brain-derived neurotrophic factor, insulin-like growth factor), oxidative damage (e.g. 8-hydroxydeoxyguanosine, malondialdehyde), and genetic biomarkers (e.g. cystatin C, calpain-10) (141). In their conclusions, the authors highlighted that accounting for these factors can improve the accuracy of outcome prognosis. At the same time, they suggested that to this end, combining information on multiple biomarkers may be necessary, and that present findings require further support from large-scale clinical trials.

Interestingly, the authors of the other, in this case systematic review (including 66 papers), concluded that there was no convincing evidence to indicate the prognostic value of genetic or biochemical markers, with considerable inconsistencies in results across selected studies (142). Their findings did, however, support the relevance of cerebral atrophy to the prognosis of post-stroke cognitive function, as previously reported by Pendlebury and Rothwell (23).

2.6.2 Predictors of post-stroke delirium

Research findings suggest that there is some overlap between prognostic factors for delirium and those identified as relevant to cognitive impairment following stroke, despite considerable differences in nature and clinical course of the two types of disorders. This perhaps seems unsurprising, given that both arise from a background of neuropathophysiology and are interrelated, although the nature of this relationship remains poorly understood (143). As identified in a recent scoping review of 25 publications (144), examples of shared predictors for post-stroke delirium and dementia that were reported across multiple studies included: older age, atrial fibrillation, previous stroke, prior cognitive decline, leukoaraiosis, cerebral atrophy, and left hemisphere and haemorrhagic index strokes.

Other risk factors seemed specific to delirium, although in some cases it is difficult to discern whether this reflects a genuine lack of associations with post-stroke cognitive impairment, or whether these variables have been rarely considered as candidate predictors for the latter. These factors included: metabolic disturbances (e.g. abnormal levels of sodium, glucose, urea nitrogen, and capillary oxygen saturation), high total number of medications, anticholinergic medications, and acute deficits and complications, namely, dysphagia, visuospatial neglect, and chest and urinary tract infections.

2.6.3 Reflection on the current stage of research

Reviews such as the ones I describe above (for other examples see: (39, 145-147)) seem now invaluable for researchers and clinicians with an interest in prognostic factors for post-stroke cognitive outcomes. Of course, as in any research area, combining findings from multiple studies provides a higher level of evidence (148, 149). However, in addition to this argument, an attempt to draw conclusions on an individual basis from the full scope of existing publications on this topic could quickly become overwhelming.

In contrast, studies on post-stroke cognition within the next prognosis theme – relating to prognostic models - appear much less evident in the current research landscape. A 2017 review described three prognostic rules for prediction of post-

stroke cognitive impairment, two of which specifically focused on development of dementia (150). However, in keeping with the narrative nature of the publication, there was no explicit mention of how the existing literature was searched, or why these particular development studies were selected. Moreover, post-stroke delirium was not within the scope of the review. As follows, the stage of progress in prognostic model research for post-stroke cognitive disorders seems at present unclear. I address this gap in the following chapter.

2.7 Summary

The ultimate purpose of prognosis research is to improve future health outcomes. Contributing to achieving this goal is a wide scope of investigations, representing four main themes: fundamental prognosis research, prognostic factor research, prognostic model research, and stratified medicine research. Many studies on post-stroke cognition have been conducted within the first two of these themes, providing insight into the natural history of cognitive change, the baseline risks of developing cognitive disorders, and factors that are associated with the latter. Currently, much less evident are examples of research pertaining to the third theme, for which the first objective is to combine information on multiple prognostic factors to quantify individual prognosis.

Chapter 3 Prognostic rules for predicting cognitive outcome following stroke: A systematic review

As I presented across the two previous chapters, to date, many studies have been conducted on post-stroke cognition within the first two prognosis research themes, as distinguished in PROGRESS – fundamental prognosis and prognostic factor research. The purpose of this chapter is to gain a better insight into current advances within the subsequent theme, relating to prognostic model research. To this end, I conducted a systematic review of prognostic rules for predicting cognitive impairment and delirium following stroke. My findings made an important contribution to informing the aims and design of subsequent studies, included as part of this thesis. This chapter is an adaptation of my published work (151).

3.1 Introduction

Increasing global prevalence and immense personal and societal costs of acquired cognitive disorders have led policymakers, researchers, and clinicians to prioritise identification of individuals at high risk. As a result, many prognostic rules for predicting cognitive impairment and decline have been developed in the general population, with a recent systematic review on this topic having identified 61 (152). The authors found the following predictors to be included across several rules: age, sex, education, physical activity, alcohol intake, body mass index (BMI), diabetes, systolic blood pressure, cholesterol levels, cardiovascular disease, depression, apolipoprotein E gene (APOE) status, and baseline performance on cognitive tasks.

Nearly half as many (27) prognostic rules have also been reported for prediction of delirium in older adult inpatients (aged over 60) (153). Across models that had been externally validated, the most frequently incorporated predictors were: age, pre-existing cognitive impairment, sensory impairment, functional disability, and severity of the acute illness.

Considering the predictors listed above, it seems that many of these variables are likely to remain relevant to the development of cognitive disorders following stroke. Nonetheless, there are strong arguments for creating dedicated prognostic rules for use in this clinical population. These include: different prevalence rates of common risk/protective factors; the importance of unique predictors, specific to the index stroke and consequent treatment; and practical considerations relevant to the acute setting, where some data may be difficult to collect (e.g. where involving an extensive assessment of function), while other information becomes easily accessible (e.g. through conducting routine blood tests). Results from a recent study provide further support for this notion (154). Specifically, the authors found that applied in a sample of nearly 1300 stroke survivors, the discriminatory power of three externally validated prognostic rules developed for use in the general population was poor (C-statistic ranging from 0.53 to 0.66).

Through scoping the literature, I concluded that there appeared to be no published systematic review addressing tools for individual cognitive outcome prognosis following stroke. To fill this gap, assisted by other researchers, I identified, described, and appraised existing prognostic rules for predicting post-stroke cognitive impairment and delirium (152). In assessing rule performance and utility, I considered the development process, and any external validation and impact studies.

3.2 Methods

This review is based on a pre-registered protocol, available on the PROSPERO International Prospective Register of Systematic Reviews (registration number: CRD42020170428). In its design, conduct and reporting, I followed the Checklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS) (155) and Preferred Reporting for Systematic Review and Meta-Analyses (PRISMA) guidelines (156). The completed PRISMA checklist is presented in Appendix 1. Together with another researcher (KM), trained and experienced in conducting systematic reviews, we independently completed all aspects of study selection, data extraction and critical appraisal. We discussed and resolved disagreements through consensus. Where an agreement could not be reached, we consulted a third, senior researcher (TQ).

3.2.1 Search strategy

We searched four electronic databases from inception to November 13, 2019: MEDLINE (via OVID), EMBASE (via OVID), PsycINFO (via EBSCO), and CINAHL (via EBSCO). I developed the search strategy (presented in full in Appendix 2) in consultation with a Cochrane Information Specialist, based on validated search filters (157-160). For all databases, the search involved terms relevant to stroke, cognition and prognosis, combined with the Boolean operator AND. I applied limits to retrieve only human studies published in English. To complement the database search, we additionally screened reference lists of relevant reviews. Based on publications identified through both procedures, we conducted backward and forward citation searches, the latter using Google Scholar.

3.2.2 Study selection

We screened titles and abstracts using the Rayyan Qatar Computing Research Institute online application (161). We applied intentionally broad inclusion criteria, aiming to identify publications reporting on the development, validation or impact assessment of prognostic rules for any global post-stroke cognitive outcome.

We included full papers, published in peer-reviewed journals. Eligible development studies recruited adult participants with a clinical diagnosis of ischaemic or haemorrhagic stroke. In relation to design, we included prospective cohort, retrospective cohort and case-control studies. Studies that self-identified as cross-sectional were eligible if predictor data related to an earlier time-point than the outcome, e.g. with information on input variables extracted from medical records. Randomised control trials were considered for inclusion if a prognostic model had been developed in the control arm or the effect of the intervention was accounted for. Regarding outcomes, we included studies applying one or more of the following assessment methods: (i) validated brief screening tools; (ii) neuropsychological batteries; (iii) expert individual or consensus diagnosis, using recognised medical classification criteria.

We excluded studies involving survivors of subarachnoid haemorrhage, due to differing pathophysiology, clinical course, and risk of unfavourable outcomes.

Further, we excluded case studies, as these could not fulfil data and analysis requirements for derivation of prognostic models. We considered rules designed to predict outcome in one specific cognitive domain (e.g. language or spatial attention) to be beyond the scope of this review. We applied no limits based on study setting, follow-up duration, or type of incorporated predictors. In the final stage, we excluded publications that did not provide a method for estimating individual outcome probability (e.g. using a mathematical formula, graphical tool or online calculator).

In relation to validation and impact studies, we applied only two inclusion criteria. Firstly, we required the availability of a published paper describing the development of the considered prognostic rule. Secondly, the study sample needed to be comparable to the derivation cohort (i.e. a population of adult stroke survivors).

3.2.3 Data extraction

We used a pre-specified, piloted proforma to extract data from selected prognostic rule development studies, including information on: study setting, development sample characteristics, predictor and outcome variables, methods of model derivation and validation, and measures of prediction rule performance. The latter was also of primary interest in relation to external validation studies. Where relevant information had not been reported, yet may have been assumed to be easily available (e.g. regarding study setting), we contacted the study authors, requesting additional details.

For recording prognostic rule validation strategies, we distinguished four levels (as I described in sections 2.4.1.1 and 2.4.2 of Chapter 2), listed in order of increasing stringency (119):

1. apparent validation – predictive ability is assessed directly in the derivation cohort;

2. internal validation – the initial dataset is split or data re-use techniques are applied, such as cross-validation or bootstrapping, to quantify overfitting and adjust for optimism in estimates of model performance;

3. temporal validation – performance is evaluated in a sample of participants recruited subsequently from the same centre(s), independently of the original data;

4. external validation – predictive ability is assessed in new and independent data, collected from an appropriate participant population in a different centre, sometimes also by different investigators.

Among measures of performance, we prioritised estimates of discrimination and calibration, as properties that are necessary (although not sufficient) to ensure practical value of prognostic tools (109). To aid interpretation of reported estimates of discriminatory power as reflected by AUROC values, we applied the following rule of thumb: <0.51 – of no value/equivalent to chance; 0.51 to 0.69 – poor; 0.70 to 0.79 – fair; 0.80 to 0.89 – good; 0.90 to 0.99 – excellent; 1.00 – perfect (162).

In cases where assessment of discrimination and/or calibration was not reported, we sought information on any alternative measures of prognostic rule performance. This particularly involved classification measures, such as: sensitivity, specificity, positive predictive value and negative predictive value.

3.2.4 Quality assessment

We assessed risk of bias for each included study using the Prediction model Risk of Bias Assessment Tool (PROBAST) (95). The tool comprises four domains: participants, predictors, outcome and analysis. Domains are appraised separately and then considered jointly to make an overall judgement. Overall risk of bias is concluded to be high if rated as high for at least one domain. PROBAST additionally incorporates an assessment of study applicability - three domains (with the exclusion of analysis) are judged based on their relevance to the population and settings targeted by the review. I had previously described some of the key considerations involved in assessing quality of prognostic studies in a focused review of rules for predicting post-stroke functional outcomes (163).

3.3 Results

Following deduplication, we initially screened 16,828 titles and abstracts (Figure 3-1). From publications considered in full-text review, we included 10 studies. We identified no additional papers through backward or forward citation searching. All relevant studies presented the development of a prognostic rule (two alternative rules in one case), with only one including a report on an external validation. As follows, we found no independent external validation publications, or studies quantifying the impact of using a prognostic rule in practice.

In total, 3143 participants from seven different Asian and European countries were involved in the development of identified prognostic rules. On average, the rules consisted of five input variables (range: 3 – 7). Predicted post-stroke outcomes included any form of global cognitive impairment, dementia and delirium. For all studies, we rated the overall risk of bias to be high.

Due to differences in clinical course, considered risk factors, and in turn – related modelling challenges – I have described prognostic rules for cognitive impairment and delirium separately. Features of identified studies are summarised in Table 3-1, while Table 3-2 presents characteristics of participant samples. Table 3-3 includes information on properties of the 11 prognostic rules, with Table 3-4 providing a general overview of types of incorporated predictors. Table 3-5 presents risk of bias ratings, using a “traffic light” colour code. Completed PROBAST forms for each study are provided in Appendix 3.

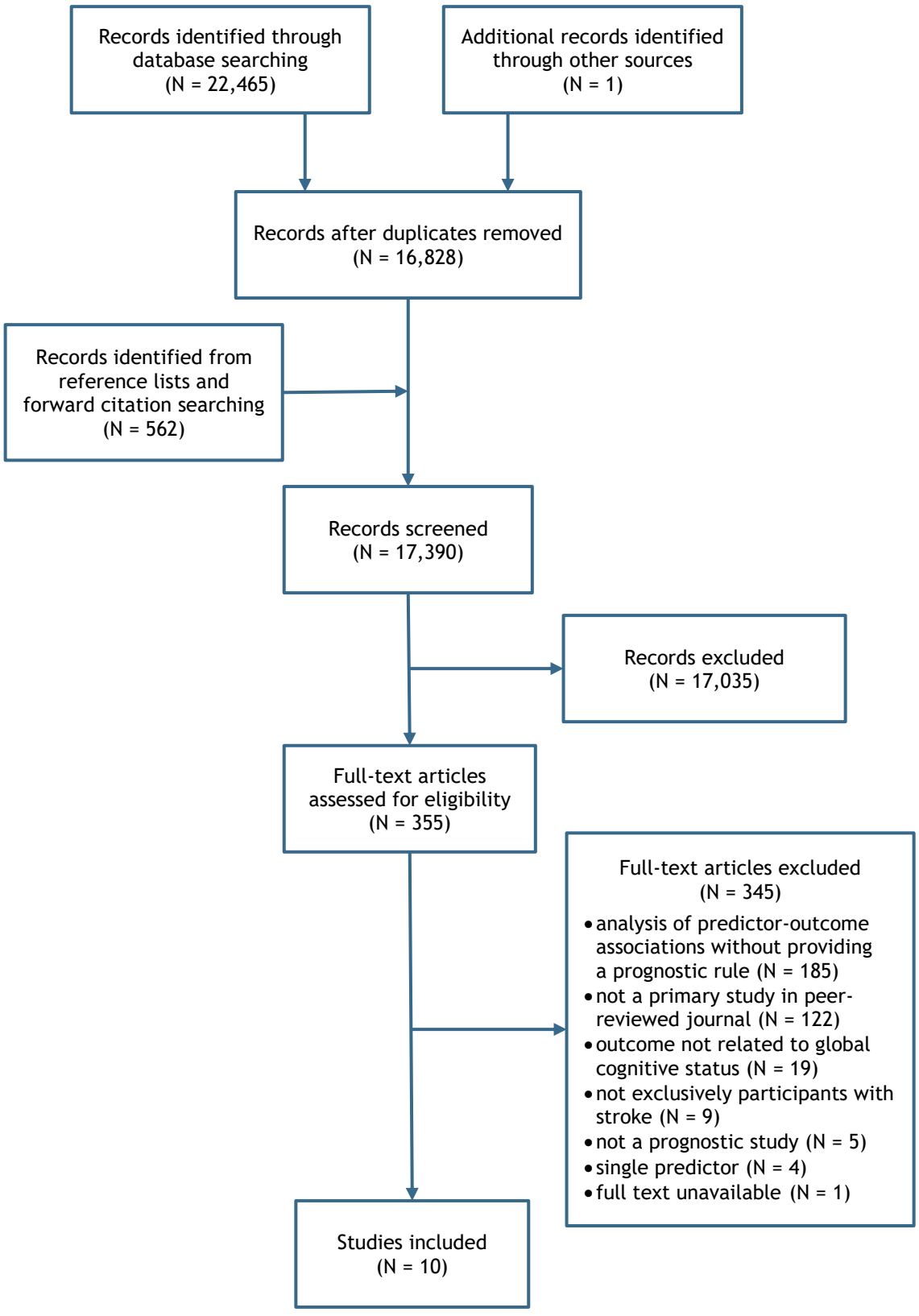


Figure 3-1 Flowchart of study selection and inclusion process.

3.3.1 Prognostic rules for cognitive impairment

3.3.1.1 Overview

Seven studies addressed the probability of developing post-stroke cognitive impairment (164-170), including one aiming to predict a favourable outcome (no cognitive impairment) (169), and two specifically focusing on risk of dementia (168, 170). The CHANGE (164) score was developed to overcome limitations of an earlier prognostic rule – $SIGNAL_2$ (167) – and was derived using the same dataset. Both rules were intended for use in cases of non-disabling stroke (modified Rankin Scale [mRS] score ≤ 2) (171). The nomogram created by Gong et al. (166) was the only tool aiming to predict cognitive impairment following intracranial haemorrhage, exclusively.

Six prediction models were derived based on logistic regression analysis, while one study applied a Chi-square Automatic Interaction Detection (CHAID) algorithm (170). On average, the identified prognostic rules included five variables (range: 3 – 7), pertaining to: demographics (in five rules), imaging findings (in five), symptom severity (in three), stroke type (in two), baseline function (in 2), and medical history (in two). Among the five studies that estimated discrimination, the reported AUROC in apparent validation ranged from 0.81 (169) to 0.91 (166). CHANGE (164) was the only prognostic rule to be externally validated, using data from a cohort of Chinese participants with ischaemic stroke (derivation cohort was from Singapore). Here, discriminatory power was found to be fair (0.75), compared to good (0.82) in apparent validation, although estimated 95% confidence intervals (CIs) overlapped.

Three studies provided a graphical assessment of calibration. For both CHANGE (164) and $SIGNAL_2$ (167), the fit between observed and predicted probabilities of cognitive impairment seemed close, with a more pronounced mismatch for the highest score values. Importantly, however, calibration of $SIGNAL_2$ (167) had only been assessed in the derivation cohort. Gong et al. (166) concluded their nomogram had good calibration in the development dataset, with best fit evident for lowest and highest scores. However, the already visible differences between observed and predicted probabilities for middle-range scores became strongly apparent upon internal validation. For lower middle-range nomogram

values, the probability of cognitive impairment was underestimated, while for higher middle-range values – considerably overestimated.

3.3.1.2 Risk of bias and applicability

For each study, we rated the risk of bias to be high in the domain of analysis. Two reasons were applicable to all cases: inappropriate handling of missing data (participants with missing data excluded/no explicit mention of approach), and not accounting for data complexities (use of analysis methods that do not allow for inclusion of censored participants). Two studies did not estimate discriminatory power (168, 170), while calibration was not assessed appropriately (165) or at all (168-170) in four studies. Assessment of rule performance was limited to apparent validation in two studies (165, 168), while no validation procedure was reported by Salihovic et al. (170).

Given the broad review question, applicability was overall of low concern, with one exception in the domain of predictors. Munsch et al. (169) obtained information on one of the input variables – stroke location – based on the outcome, using lesion symptom mapping, rather than prior to outcome assessment.

3.3.2 Prognostic rules for delirium

3.3.2.1 Overview

Three studies aimed to predict risk of post-stroke delirium (172-174), producing four prognostic rules – in a pilot attempt (recognizing sample size limitations), Kostalova et al. (172) presented two alternatives. All prognostic models were developed based on logistic regression analysis. On average, the prediction rules included five variables (range: 4 – 7), pertaining to: demographics (in all rules), imaging findings (in two), symptom severity (in two), stroke type (in three), baseline function (in one), acute medical complications (in two), and laboratory markers (in two) (Table 3-3 and Table 3-4). The latter two variable categories were unique to prognostic rules for delirium, not being included among predictors of cognitive impairment.

On account of the fluctuating course of delirium, in all studies the outcome was assessed on multiple occasions. Kostalova et al. (172) and Kotfis et al. (173) conducted assessments daily for up to eight and six days, respectively, including the day of hospital admission. Oldenbeuving et al. (174) screened for delirium on two separate days within a seven-day period from admission.

Out of the two studies that assessed discrimination in apparent validation, Oldenbeuving et al. (174) reported the higher estimate – AUROC = 0.84, compared to 0.80 and 0.73 reported by Kotfis et al. (173), for outcome measured at an earlier and later time-point, respectively. Oldenbeuving et al. (174) had also applied the most stringent form of validation – temporal – reporting an AUROC of 0.83.

3.3.2.2 Risk of bias and applicability

For all studies, we rated the risk of bias to be high in domains of outcome and analysis. Regarding the former, common concerns related to lack of blinding to predictors, or even use of predictor knowledge to inform outcome assessment. In terms of analysis, we judged the risk of bias to be high due to insufficient sample size, inappropriate handling of missing data and/or data complexities, and no evaluation of rule calibration. Assessment of discrimination was omitted from the study by Kostalova et al. (172), while Kotfis et al. (173) applied no method to adjust for optimism in estimating the performance of DELIAS.

We rated applicability to be of high concern in studies by Kostalova et al. (172) and Oldenbeuving et al. (174), due to risk of overlap in timing of predictor and outcome assessments. Based on reported information, we were not able to ascertain whether a similar issue applied to the study by Kotfis et al. (173), therefore in this case we considered applicability to be unclear.

Table 3-1 Characteristics of included studies.

Study	Country	Setting	Recruitment period	Design	Follow-up duration	Stroke type	Exclusion criteria of note
Prognostic rules for cognitive impairment							
Chander et al., 2017; CHANGE	Singapore	Tertiary outpatient stroke clinic	Jan 2008 to Dec 2012	Retrospective cohort	3 to 6 months	Ischaemic	Discharge mRS > 2; pre-stroke cognitive impairment; neurologic or psychiatric comorbidities; impairment impeding cognitive assessment
Ding et al., 2019	China	Neurology department of university hospital	June 2017 to May 2018	Prospective cohort	6 to 12 months	Ischaemic	Major mental illness; pre-existing dementia; impairment impeding cognitive assessment
Gong et al., 2019	China	Hospital rehabilitation department	Jan 2016 to Oct 2018	Retrospective cohort	3 to 6 months	Supratentorial haemorrhage	Pre-existing dementia; previous stroke
Kandiah et al., 2016; $SIGNAL_2$	Singapore	Tertiary outpatient stroke clinic	Jan 2008 to Dec 2012	Retrospective cohort	3 to 6 months	Ischaemic	Discharge mRS > 2; pre-stroke cognitive impairment; neurologic or psychiatric comorbidities; impairment impeding cognitive assessment
Lin et al., 2003	Taiwan	Neurology department of university hospital	Nov 1995 to Oct 1999	Prospective cohort	3 months	Ischaemic	Severe medical comorbidity; pre-existing dementia with nonvascular aetiology

Table 3-1 Characteristics of included studies. *Continued*

Study	Country	Setting	Recruitment period	Design	Follow-up duration	Stroke type	Exclusion criteria of note
Prognostic rules for cognitive impairment							
Munsch et al., 2016	France	Neurology department of university hospital	Jun 2012 to Feb 2015	Prospective cohort	3 months	Supratentorial ischaemia	History of cerebral infarct with functional deficit; pre-existing psychiatric disorders other than depression; pre-existing dementia
Salihovic et al., 2018	Bosnia and Herzegovina	Neurology department at a university clinical centre	Sep 2011 to Aug 2012	Prospective cohort	12 months	Ischaemic and haemorrhagic	Pre-existing cognitive impairment, recurrent stroke, aphasia impeding cognitive assessment
Prognostic rules for delirium							
Kostalova et al., 2012	Czech Republic	Stroke unit of university hospital	Jan 2009 to Mar 2010	Prospective cohort	7 days	Ischaemic and haemorrhagic	History of head trauma, neurosurgery or psychosis; RASS <-3 (deep sedation, unarousable)
Kotfis et al., 2019; DELIAS	Poland	Neurology department of district general hospital	Jun 2015 to Mar 2018	Prospective cohort	5 days	Ischaemic	Haematology disorders
Oldenbeuving et al., 2014	Netherlands	Stroke units of two general hospitals	1-year period	Prospective cohort	Up to 7 days	Ischaemic and haemorrhagic	Severe intellectual disability; severe language barrier

mRS indicates modified Rankin Scale; RASS, Richmond Agitation-Sedation Scale.

Table 3-2 Participant characteristics for included studies.

Study	Model development sample, N	Age, years, mean (SD)	Women, N (%)	NIHSS Score, median (IQR)	Participants with outcome, N (%)
Prognostic rules for cognitive impairment					
Chander et al., 2017; CHANGE	209	61.7 (12.5)	67 (32.1%)	Not reported	78 (37.3%)
Ding et al., 2019	145	No cognitive disorder group: Mdn = 61, IQR: 48.5 - 69.0; cognitive disorder group: Mdn = 64, IQR: 60.0 - 73.0	42 (29.0%)	No cognitive disorder group: 3.0 (1.0 - 5.0); cognitive disorder group: 4.0 (2.0 - 7.0)	77 (53.1%)
Gong et al., 2019	92	57.3 (12.2)	28 (30.4%)	Not reported	69 (54.3%)/127*
Kandiah et al., 2016; $SIGNAL_2$	209	61.7 (12.5)	67 (32.1%)	Not reported	78 (37.3%)
Lin et al., 2003	283	64.4 (8.4)	95 (33.6%)	M = 3.6, SD = 3.1	26 (9.2%)
Munsch et al., 2016	198	No cognitive disorder group: Mdn = 60; range: 29 – 84; cognitive disorder group: Mdn = 69, range: 34 – 95	77 (35.8%)/215**	No cognitive disorder group: 3.0, range: 1.0 – 10.0; cognitive disorder group: 4.0, range: 1.0 – 25.0	77 (38.9%)
Salihovic et al., 2018	275	Females: 66.3 (2.0); Males: 65.1 (1.5)	103 (37.5%)	Score of 0 - 7: N = 163 (59.3%); score of 8 - 14: N = 89 (32.4%); score > 14: N = 23 (8.4%)	190 (69.1%)

Table 3-2 Participant characteristics for included studies. *Continued*

Study	Model development sample, N	Age, years, mean (SD)	Women, N (%)	NIHSS Score, median (IQR)	Participants with outcome, N (%)
Prognostic rules for delirium					
Kostalova et al., 2012	100	73.5 (11.5)	47 (47.0%)	No cognitive disorder group: 9.0, 5th - 95th percentile range: 4.0 – 17.0; cognitive disorder group: 11.0, 5th - 95th percentile range: 5.0 - 16.0	43 (43.0%)
Kotfis et al., 2019; DELIAS	1001	Mdn = 71.0, IQR: 64.0 - 82.0	478 (47.8%)	No cognitive disorder group: 8.0 (4.0 - 14.0); cognitive disorder group: 18.0 (12.0 - 21.5)	172 (17.2%)
Oldenbeuving et al., 2014	527	72.0, range: 29.0 – 96.0	239 (45.4%)	5.0, range: 0.0 – 36.0	62 (11.8%)

*Combined development and validation cohorts.

**Sample before excluding subjects with no outcome data.

IQR indicates interquartile range; M, mean; Mdn, median; NIHSS, National Institutes of Health Stroke Scale; SD, standard deviation.

Table 3-3 Characteristics of included prognostic rules.

Study	Predictors	Outcome	Assessment	Validation, strategy and corresponding AUROC (95% CI)
Prognostic rules for cognitive impairment				
Chander et al., 2017; CHANGE	Age, education, acute nonlacunar cortical infarcts, chronic lacunes, white matter hyperintensities, global cortical atrophy	Cognitive impairment	Structured clinical interview and MMSE, MoCA if further confirmation required	Apparent: 0.82 (0.76, 0.88); temporal: 0.78 (0.71, 0.85); external: 0.75 (0.71, 0.79)
Ding et al., 2019	Age, education, acute nonlacunar infarcts, periventricular hyperintensity, diabetes mellitus	Cognitive impairment	MMSE, MoCA, neuropsychological battery, assessment based on CDR and DSM-4 criteria	Apparent: 0.88 (0.83, 0.94)
Gong et al., 2019	Intraventricular hemorrhage, GCS score, bleeding volume	Cognitive impairment	MMSE	Apparent: 0.91; internal with data splitting based on recruitment period: 0.92 (CIs not reported)
Kandiah et al., 2016; $SIGNAL_2$	Age, education, acute nonlacunar cortical infarcts, chronic lacunes, white matter hyperintensities, global cortical atrophy, intracranial stenosis	Cognitive impairment	Structured clinical interview and MMSE, MoCA if further confirmation required	Apparent: 0.83 (0.77, 0.88); temporal: 0.78 (0.70, 0.85)
Lin et al., 2003	Age, occupation, previous stroke, vascular territory of infarction, NIHSS score, MMSE score, FIM motor score	Dementia	Consensus diagnosis based on CDR CERAD, neuropsychological battery, and criteria of Alzheimer's Disease and Related Disorders Association, ICD-10NA, NINDS, and NINDS-AIREN	Apparent: 93.4% of participants correctly classified according to outcome

Table 3-3 Characteristics of included prognostic rules. ***Continued***

Study	Predictors	Outcome	Assessment	Validation, strategy and corresponding AUROC (95% CI)
Prognostic rules for cognitive impairment				
Munsch et al., 2016	Age, infarct volume, NIHSS score, stroke location expressed as number of eloquent voxels from voxel-based lesion-symptom mapping maps	Good cognitive outcome (no cognitive impairment)	MoCA	Apparent: 0.81 (0.75, 0.87); internal with 10-fold cross validation and 1000 bootstrap replications: 0.77 (0.69, 0.84); internal with data splitting based on recruitment period: 0.78 (0.70, 0.85)
Salihovic et al., 2018	Complex figure test score, narrative memory score, numerical memory score	Vascular dementia	Diagnosis using clinical exams and neuropsychological testing, based on DSM-4 and ICD-10	Not assessed
Prognostic rules for delirium				
Kostalova et al., 2012; Rule 1	Age, intracerebral haemorrhage, lesion volume, gamma-glutamyl transferase, bilirubin	Delirium	Consensus diagnosis based on DSM-4 criteria, CAM-ICU	Internal with 2-fold cross-validation; correctly classified 69.0% of subjects with delirium and 84.2% without
Kostalova et al., 2012; Rule 2	Age, intracerebral haemorrhage, lesion volume, SOFA-Max	Delirium	Consensus diagnosis based on DSM-4 criteria, CAM-ICU	Internal with 2-fold cross-validation; correctly classified 65.1% of subjects with delirium and 80.7% without

Table 3-3 Characteristics of included prognostic rules. ***Continued***

Study	Predictors	Outcome	Assessment	Validation, strategy and corresponding AUROC (95% CI)
Prognostic rules for delirium				
Kotfis et al., 2019; DELIAS	Age, NIHSS score, hemianopia, aphasia, neutrophil to lymphocyte ratio, leukocytes, c-reactive protein	Early-onset delirium (up to 24 hours), delirium up to 5 days	CAM-ICU and investigator assessment based on DSM-5 criteria	Apparent, for early onset delirium: 0.80; for delirium up to 5 days: 0.73 (CIs not reported)
Oldenbeuving et al., 2014	Age, stroke subtype, NIHSS score, infection	Delirium	CAM	Apparent: 0.84 (0.80, 0.89); temporal: 0.83 (0.77, 0.90)

AUROC, area under the receiver operating characteristic; CAM(ICU), Confusion Assessment Method (for the Intensive Care Unit); CDR, Clinical Dementia Rating Scale; CERAD, Consortium to Establish a Registry for Alzheimer's Disease; CI, confidence interval; DSM, Diagnostic and Statistical Manual of Mental Disorders; FIM, Functional Independence Measure; GCS, Glasgow Coma Scale; ICD-10NA, International Classification of Diseases, tenth revision: Neurological Adaptation; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; NIHSS, National Institutes of Health Stroke Scale; NINDS(AIREN), National Institute of Neurological Disorders and Stroke (Association Internationale pour la Recherche et l'Enseignement en Neurosciences); SOFA-Max: Sequential Organ Failure Assessment, maximum score.

Table 3-4 Types of variables included in prognostic rules.

Study	Demographics	Medical history	Symptom severity	Stroke type	Imaging findings	Acute medical complications	Laboratory markers	Baseline function
Prognostic rules for cognitive impairment								
Chander et al., 2017; CHANGE	✓				✓			
Ding et al., 2019	✓	✓			✓			
Gong et al., 2019			✓	✓	✓			
Kandiah et al., 2016; $SIGNAL_2$	✓				✓			
Lin et al., 2003	✓	✓	✓	✓				✓
Munsch et al., 2016	✓		✓		✓			
Salihovic et al., 2018								✓
Prognostic rules for delirium								
Kostalova et al., 2012; Rule 1	✓			✓	✓		✓	
Kostalova et al., 2012; Rule 2	✓			✓	✓	✓		
Kotfis et al., 2019; DELIAS	✓		✓				✓	✓
Oldenbeuving et al., 2014	✓		✓	✓		✓		

Table 3-5 Risk of bias rating for included studies.

Study	Participants	Predictors	Outcome	Analysis	Overall rating
Prognostic rules for cognitive impairment					
Chander et al., 2017; CHANGE					
Ding et al., 2019					
Gong et al., 2019					
Kandiah et al., 2016; $SIGNAL_2$					
Lin et al., 2003					
Munsch et al., 2016					
Salihovic et al., 2018					
Prognostic rules for delirium					
Kostalova et al., 2012					
Kotfis et al., 2019; DELIAS					
Oldenbeuving et al., 2014					

Green indicates low risk of bias; amber – unclear risk of bias, red – high risk of bias.

3.4 Discussion

We identified 11 prognostic rules for prediction of post-stroke cognitive outcomes, three of which had been published just within a year before the literature search date. However, we found no independent external validation studies, or reports on assessing implementation of prognostic rules in practice. Research into prognosis of post-stroke cognitive outcomes is an expanding area, but still at its early stages, with a primary focus on development of novel strategies, rather than validation or application.

3.4.1 Clinical implications

Based on our findings, I cannot indicate preferred prognostic rules for prediction of either post-stroke delirium or longer-term cognitive outcome. All included studies had strengths and limitations. The highest discriminatory power in apparent validation was reported by Gong et al., whose study was also the only one to be rated as having low risk of bias in as many as three out of four domains. However, both development and validation cohorts were small, and graphical assessment revealed considerable rule miscalibration.

Chander et al. (CHANGE) and Oldenbeuving et al. applied the most stringent validation strategies out of studies predicting post-stroke cognitive impairment and delirium, respectively. Both publications, moreover, had the advantage of producing clear scoring systems, allowing easy estimation of individual prognosis. Conversely, the same two studies had the highest number of domains rated as high risk of bias (three out of four). Although it is important to highlight that CHANGE was the only externally validated prognostic rule, it is nonetheless arguable whether the tool's reported fair discriminatory power would be considered sufficient to merit implementation in clinical practice.

A fundamental challenge is that without external validation studies, the generalisability of developed prognostic rules cannot be assessed, or their predictive accuracy directly compared against one another. Further, to choose an optimal prognostic rule, it is also essential to consider the target population and setting. Many tools may not be applicable in an unselected stroke population, for example, where development cohorts exclusively comprised

survivors of ischaemic stroke (164, 165, 167-169, 173), intracranial haemorrhage (166), or non-disabling stroke (164, 167).

Generalisability may also be compromised if a prognostic rule relies on predictor information that is not routinely available in all healthcare systems. An example is incorporation of scores based on MRI scans, which may not be attainable in resource-poor settings. Apart from considerations relevant from clinician perspective, it is also necessary to account for concerns raised by stroke survivors and their families, determining at the individual level whether the process of estimating future cognitive outcomes and its consequences is acceptable (175).

3.4.2 Research challenges

The processes of data extraction and quality assessment have highlighted some of the many challenges inherent to prognostic research. Collecting longitudinal data from large study samples is associated with high resource requirements and a prolonged delay from project inception to producing first research outputs. As such, it is not surprising that many investigators opt for use of existing datasets. The practical advantages of this approach, however, often come at a cost of obtaining relevant data. For example, given that Chander et al. (164) reported NIHSS scores for the external validation cohort, but not for the development cohort, it seems this information was not available for the latter. Consequently, despite existing evidence on the importance of this variable, stroke severity could not have been considered as a predictor for inclusion in SIGNAL$_2$ (167) or CHANGE (164).

An additional trade-off occurs with use of records from routine care registries. In this case, access to data from large clinical populations, unaffected by participation bias, is coupled with quality concerns. The encompassed individual-level information is not collected for research purposes, and therefore it cannot be expected that variables are measured consistently, in adherence to standardised protocols, such as used in research studies (176, 177).

A number of challenges are also specific to research into stroke and cognition. Participant deaths, high rates of losses to follow-up, and incomplete assessments

due to stroke-related impairments (74) can all contribute to missing data and biased study samples, where healthier individuals are overrepresented. Moreover, there is no consensus method of diagnosing post-stroke cognitive disorders (178). Adhering to medical classification criteria offers the most holistic approach to assessing cognitive function, with information from multiple sources being considered, however it also introduces incorporation bias. For example, a diagnosis of delirium according to DSM-5 criteria (6), requires obtaining evidence of a potential cause, such as infection. In other words, knowledge of predictor status is used to inform outcome assessment. Consequently, the strength of the association between the two can be overestimated, in turn leading to optimistic estimates of prognostic rule performance.

The latter may also result from recruiting participants who experienced cognitive disorders prior to index stroke (116). The included studies that aimed to predict cognitive impairment had indeed avoided this issue through applied exclusion criteria. However, this selection process leads to obtaining a study sample that is not representative of a real-world stroke population. Given the overlap between risk factors for stroke and age-related cognitive decline, it is not surprising that the prevalence of pre-stroke dementia is estimated to be around twice as high (10%) as in the general population (20, 23), with even more stroke survivors likely to have experienced milder forms of pre-stroke cognitive impairment.

Many of the described pitfalls seem impossible to avoid. However, through our risk of bias assessment, we also identified methodological problems that could have been at least partially ameliorated. For example, the most common approach we observed to model development involved use of logistic regression in a complete-case analysis. The consequent exclusion of participants could have been avoided through applying data imputation techniques and conducting a time-to-event analysis, retaining those with incomplete follow-up (116).

Another recurring problem relates to relying on significance in univariable analysis for selection of input variables, which may lead to omission of important predictors (179). For this task, use of a nonstatistical strategy is recommended, where predictors are chosen based on previous evidence, and in view of

feasibility and measurement reliability. Recognition of such amendable limitations, and application of the comprehensive, rigorous and explicit guidance presented within the recently published PROBAST tool, can help raise standards in design, conduct and reporting of future prognostic studies.

3.4.3 Strengths and limitations

To my knowledge, this is the first systematic review to focus on prognostic rules for prediction of post-stroke cognitive outcome. The opportunity to use the relatively novel PROBAST tool posed an important advantage to completing this work. Lack of a consensus approach to risk of bias assessment has limited previous reviews in the area of prognostic research (180). A further strength relates to tailoring the search strategy, and applying broad inclusion criteria, so as to promote comprehensiveness.

However, due to limited resources, we only retrieved studies published in English. Moreover, requiring that publications provide a method to estimate the individual probability of cognitive outcomes, we would have excluded studies relying on more complex prediction techniques, such as machine learning. However, this was deliberate to ensure usefulness of my review to clinicians and researchers, through focusing on methods which allow immediate application, provided predictor information is readily available.

3.4.4 Future directions and specific thesis aims

This systematic review of prognostic rules (including development procedures and content) has complemented the literature review presented in the Introduction and allowed me to identify areas of research that so far seem to have received little attention. Firstly, modifiable risk and protective factors, e.g. pertaining to lifestyle, are rarely considered among predictors of post-stroke cognitive outcome. Their relevance to cognition has been evidenced by prognostic model studies in the general population (152), and it seems implausible that such associations would cease to be important following stroke. In addition to the possibility of explaining some of the variability in post-stroke cognitive function, the importance in identifying modifiable predictors lies in representing an opportunity to reduce the risk of an unfavourable outcome.

The second topic that seems under-researched, relates to circumstances in which particular variables influence post-stroke cognitive outcome. As evident, for example, in cases of multimorbidity, relevant factors can appear in combination rather than isolation. Patterns of co-occurrence may determine how cognition is ultimately affected. Moreover, individual characteristics may be relevant to post-stroke cognition in different ways. For example, some variables may primarily contribute to pre-stroke cognitive decline, with others affecting clinical features of the index stroke, or the longer-term recovery process.

The latter point ties into a third issue – accounting for changes over time. As discussed in the Introduction, evidence suggests that cognitive function can fluctuate even in the chronic phase after stroke. With outcomes measured at a single time-point, as done in reviewed prognostic studies, it is not possible to determine whether individual cognitive status is following a stable, improving or declining trajectory. Individuals recognised as "cognitively intact" may still be experiencing meaningful deterioration in cognitive function, and therefore, in the longer-term, be at greater risk of developing dementia than those "cognitively impaired", but stable or improving.

In the following thesis chapters, I will address these three issues, with an overarching aim to improve our understanding of the cognitive change that occurs following stroke and its associations with individual characteristics. Findings from conducted studies may not only contribute to increased accuracy of future prognostic models, but also inform strategies to prevent unfavourable outcome and improve cognitive function.

3.5 Summary

Prognostic model research relevant to post-stroke cognitive outcomes is at a relatively early stage. From 11 identified prognostic rules (7 predicting cognitive impairment and 4 delirium) only one had been externally validated, and none had been assessed in terms of impact. Limited evidence regarding performance and generalisability, coupled with a high risk of bias in all model development studies, hinders recommending use of specific prognostic rules in clinical practice. Nonetheless, findings from the reviewed studies have important implications for future research, including subsequent projects in my thesis.

Chapter 4 Potentially modifiable predictors of post-stroke cognition: Are physical activity, sedentary behaviour, and social engagement relevant?

Through my review of the existing literature, I identified three themes relevant to prognosis research on post-stroke cognition that so far seem under-investigated. In this chapter, I address the first one – the relevance of potentially modifiable factors to cognitive function following stroke. Using data from a large, population-based general-purpose cohort, the UK Biobank, I conducted two separate studies in a subsample of participants with a history of stroke and TIA. In the first study, I focused on habitual physical activity and sedentary behaviour, and in the second – on different aspects of social engagement.

After defining key terms, I introduce proposed theories and evidence in support of an association between these factors and cognitive function, some of which point to the possible causal nature of these relationships. I then present a shared Methods section, following which I report and discuss the results of each study separately. Finally, I draw general conclusions from both investigations, considering the implications of my findings to the development of novel interventions for improving cognitive outcomes, as well as the applicability of such data resources as the UK Biobank for future prognostic stroke research.

4.1 Introduction

In prognosis research, if the sole priority was to create highly accurate prognostic tools, predictors would only be regarded for the strength of their independent association with an outcome. Yet, arguably, there is another crucial task to consider – understanding which of these predictors can be modified in order to improve the outcome (99). In relation to cognitive impairment and decline, two such factors – physical activity patterns and social engagement – were first targeted decades ago (e.g. see 181, 182).

Since then, research interest in physical activity as a modifiable determinant of cognitive function from childhood to older age has increased exponentially, resulting in a considerable body of supportive evidence (183, 184). In contrast, although many studies have successfully demonstrated an association between social engagement and cognition (185, 186), the presence of a causal relationship is still largely debated (187). Nonetheless, concepts of healthy cognitive aging have incorporated social engagement as one of their key components (188, 189).

4.1.1 Definitions

Physical activity is defined as “any bodily movement produced by skeletal muscles that results in energy expenditure” (190, page 126). Based on oxygen consumption, physical activity can be categorised according to three levels of intensity: light (e.g. walking, washing dishes); moderate, characterised by an increased heart rate, breathing harder, and feeling warmer than normal (e.g. mowing the lawn, dancing); and vigorous, associated with sweating and breathing hard and fast (e.g. running, carrying heavy loads) (191).

In addition to assessing one’s physical activity as an essential component of a healthy lifestyle, the importance of considering sedentary behaviour as an independent factor has been increasingly recognised (192). The term ‘sedentary behaviour’ encompasses waking-time activities that require low levels of energy expenditure, and are performed while sitting, reclining or lying down. In the context of cognition, it is also relevant to note a novel distinction being made between mentally active sedentary behaviours (e.g. doing a crossword puzzle) and mentally passive ones (e.g. watching TV) (193).

In comparison to the above, social engagement appears to be a concept more difficult to define, with apparent inconsistencies in operationalisation and measurement across studies. However, three domains of interest are typically distinguished (194): social networks, relating to structure, composition and content of an individual’s interpersonal ties (195-197); social support, relating to the level of emotional and instrumental resources available to an individual (198, 199); and social participation, relating to involvement in activities with a social component (200, 201). Another important distinction is made between

objective measures of social engagement (e.g. how often do you meet with friends) and subjective measures (e.g. feeling lonely) (202, 203).

4.1.2 Routes to affecting cognition

There are many putative pathways through which physical activity, sedentary behaviour and social engagement may influence cognitive functioning. One perspective focuses on indirect effects, driven by impacts on general health and wellbeing.

Regarding the relationship between physical activity and health, research has provided a particularly rich evidence base for the formers benefits, including: reduced blood pressure and risk of blood clot formation, enhanced insulin sensitivity and glucose tolerance, and improvements in body composition (increase in lean mass and reduction in fat), blood distribution, and plasma lipid and lipoprotein profile (204). As follows, physical activity is associated with a reduced likelihood of developing cardiovascular diseases, which are risk factors for cognitive impairment and decline.

Across multiple observational and experimental studies, findings regarding specific effects of sedentary behaviour on physiology appear overall less consistent (205); although a number of reports have indicated that prolonged sitting results in detrimental changes to insulin sensitivity, glucose tolerance and plasma triglyceride levels (206). Evidence moreover suggests that sedentary behaviour is associated with an increased risk of fatal and non-fatal cardiovascular diseases and events (207).

In addition, activity patterns may affect cognitive functioning through impact on mental health. Specifically, sedentary behaviour has been associated with depression and depressive symptoms (205), while physical activity has been found to favourably influence mood. The latter relates to both enhancing positive affect, as well as ameliorating depression, and reducing anxiety and stress (208).

Changes in stress levels have also been proposed as the primary pathway through which health is affected by factors reflecting social engagement (209).

Perceived, involuntary social isolation is hypothesised to evoke a physiological stress response, which as a chronic state may lead to dysregulation of the endocrine, immune, and cardiovascular systems (210-212). Consequently, this experience is associated with increased blood pressure and heart rate, chronic inflammation, poor sleep quality, slower wound healing and greater susceptibility to infection (213-217). These cumulative processes have been described as causing overall accelerated aging, which also involves neurodegeneration, mostly attributed to elevated levels of glucocorticoids (209, 211, 218). In addition, stress is considered to trigger unhealthy behaviours, such as excessive alcohol consumption (219), smoking (220) and disordered eating (221, 222).

Conversely, social support has been associated with improved outcomes following health and life-threating events, such as transport accidents (223), heart failure (224), and surgery (225, 226). These observations are attributed to social support acting as a buffer against stress, induced by the event, as well as enhancing adherence to medical advice (e.g. taking prescription medication).

Alongside effects obtained by modifying health and wellbeing, there is potential for activity patterns and social engagement to influence neural structure and function in a more direct manner. One concept that simultaneously accounts for the contribution of both types of factors to brain health is embedded in a line of animal model studies on the impact of an enriched environment (188). Such an environment, of which opportunity for physical activity and social interaction are key constituents, provides complex stimulation for its inhabitants. Living in these conditions is argued to enhance neuronal plasticity on molecular, cellular and structural levels, and thus benefit cognitive functioning (227). In this context, social isolation and excess of mentally passive sedentary behaviours is considered equivalent to functioning in an impoverished environment, associated with cognitive decline (228).

4.1.3 Physical activity, sedentary behaviour, social engagement, and cognition in the context of stroke

There is a clear application for the described concepts to the context of stroke – a highly stressful, potentially life-threatening event, typically associated with

prevalent cardiometabolic comorbidities. Adding to this notion are proposed further dual-path effects of environmental enrichment, specific to neurological and cognitive recovery following brain damage. The first hypothesised route, beginnings of which reach childhood, is through enhancing cognitive reserve, in turn reducing the manifestation of neuropathology and supporting functional compensation (229). The second route involves enhancement of multiple, beneficial physiological processes in the acute and subacute stroke phases, including: synaptogenesis, axonal sprouting, axonal and dendritic remodelling, neurogenesis, and angiogenesis (230).

In relation to physical activity itself, there are also some indications that its role in modifying cognitive outcomes could begin even prior to index stroke. In a study on mice, the authors reported that engaging in voluntary exercise for four weeks preceding experimental traumatic brain injury resulted in increased activation of anti-apoptotic and anti-inflammatory pathways, and improved recovery of cognitive function (231). In humans, on the other hand, an association has been observed between physical activity prior to stroke and alleviated stroke severity – an important determinant of cognitive function (232).

Although currently there seem to be no findings supporting a direct link between improved post-stroke cognition and pre-stroke physical activity, there is high quality evidence indicating a favourable effect of post-stroke physical activity. In a meta-analysis of 14 randomised controlled trials (N = 736), the authors concluded that structured physical activity training had a positive overall impact on post-stroke cognitive performance, with a small to moderate mean effect size (233). Cognitive gains were largest for programmes that combined aerobic and strength training, and were apparent even where the intervention was introduced in the chronic phase of stroke. When domain-specific performance was considered, improvement was reported for attention and processing speed, while no significant effects were observed for executive function or working memory.

These results may appear conclusive, however the topic of physical activity and post-stroke cognition is not yet exhausted. Alongside reported methodological limitations of studies included in the review, it is noteworthy that structured

training does not cover the whole spectrum of physical activities that people can engage in. This is important, as despite interventions being tailored to accommodate post-stroke motor deficits, it is likely that many stroke survivors would find objective and subjective barriers to engaging in such programmes (234, 235); particularly, as it is plausible that those who agreed to participate and completed the interventions, were more active, health-orientated individuals.

As follows, there is still a need to improve our understanding of how typical, day-to-day physical activity, encompassing chores and leisure activities alongside intentional exercise, is associated with post-stroke cognition. An apparent link could provide grounds for implementing less demanding interventions, more acceptable and sustainable in the long term for a wider population of stroke survivors. Considering physical activity and sedentary behaviour in conjunction, a clinically meaningful change could be sought even just by breaking up prolonged sitting time with brief bouts of light-intensity indoor walking (236, 237).

This notion seems of particular relevance given that stroke survivors are reported to spend significantly more time sitting than their peers (10.9 hours/day vs 8.2 hours/day) (238). However, to date, little is known about the effects of interrupting sedentary behaviour on post-stroke outcomes, beyond one study demonstrating a decrease in systolic blood pressure (239). As such, it is the cumulative evidence, pieced together from studies involving animal models and non-stroke populations (240), that is the main indicator of a possible detrimental effect of sedentary behaviour on cognitive function following stroke.

Similarly, more research is needed to describe the relationship between social life and post-stroke cognition – one that is likely complex, with social isolation recognised as a risk factor for stroke (241), and stroke contributing to social isolation (242-244). At present, it seems that only one study (N = 272) has addressed this topic directly (245). The authors found that baseline social ties and emotional support were independent predictors of better performance on a cognitive summary score at six months, while emotional support was further associated with greater improvement from baseline to follow-up. When individual tasks were considered, social ties predicted immediate and delayed

recall, while emotional support was associated with immediate recall only. Yet no significant relationships were found for tasks assessing attention, language and executive function, nor for performance on the MMSE.

Complementing these observations are reports of stroke survivors showing progressive functional improvement with high levels of social support, while (ultimately) experiencing functional deterioration with low support (246-248). Although cognition was not assessed in these studies, if an association exists between social engagement and post-stroke functional recovery, one with cognitive recovery also seems plausible.

4.1.4 The present studies

I conducted two observational studies to investigate the associations of post-stroke cognitive performance with: physical activity and sedentary behaviour (Study A), and social engagement (Study B). I made three assumptions regarding Study A: i) the effects of physical activity and sedentary behaviour are independent from one another (to a degree), and therefore both should be accounted for; ii) the effects of physical activity and sedentary behaviour may depend on their type; while also iii) the accumulated time spent being physically active and sedentary may be relevant to cognitive performance.

In relation to Study B, I assumed that both objective and subjective aspects of social engagement are relevant to post-stroke cognitive function, and their effects might differ. Further, based on the described studies in stroke populations, I hypothesised that associations for all predictors of interest are likely to depend on what cognitive function is being assessed. I aimed to ensure that plausible confounders were accounted for, recognising that many variables may be simultaneously relevant to: a) activity patterns and cognition, and b) social engagement and cognition.

4.2 Methods

I used anonymised, individual participant level data held in the UK Biobank. The UK Biobank project is overseen by the National Health Service (NHS) National Research Ethics Service (approval letter dated 17th June 2011, Ref 11/NW/0382)

and has received ethical approval from the Community Health Index Advisory Group (approved 7th December 2006, application no. 06-007). I conducted this research under application no. 17689. In reporting the two studies, I followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (249). Study B, relating to social engagement, has been published (250). I edited and expanded the content of my paper for the purpose of this thesis chapter.

4.2.1 Study setting and participants

The UK Biobank includes data from over 502,500 participants. Baseline assessments took place between 2006 and 2010, across 22 centres in the United Kingdom. During the study visit, participants answered questions regarding sociodemographic, health, mood and lifestyle factors, completed cognitive tasks, and had a range of physical measurements taken. A more detailed description of the UK Biobank has been provided previously (251).

I focused on baseline data from participants who self-reported a history of stroke or TIA (data field 20002). I included cases of TIA on account of similar prevalence of risk factors for cognitive decline as in stroke populations (252), and evidence of longer-term cognitive sequalae (253). As per UK Biobank assessment procedures, information on medical conditions was obtained during a two-stage process. Firstly, through a touch-screen questionnaire, participants were asked whether they had a history of one or more illnesses, including stroke. Responses were subsequently confirmed during a verbal interview with a trained nurse. In cases where the participant was uncertain of the type of condition they had had, they were asked to describe the illness, so that the nurse could assist in defining it. If the interview revealed an erroneous indication of a medical condition, the initial response was amended.

To identify eligible study participants, I considered four items in the repository under the cerebrovascular disease category: “stroke” (code 1081), “ischaemic stroke” (code 1583), “brain haemorrhage” (code 1491), and “transient ischaemic attack” (code 1082). To ensure a consistent approach throughout my thesis, unlike in the published version of Study B (250), I did not include cases of

subarachnoid haemorrhage. I further excluded participants who reported the occurrence of stroke or TIA before the age of 18.

4.2.2 Measures

4.2.2.1 Predictors of interest in study A

I included three self-reported measures of physical activity and three of sedentary behaviour, relating to duration on a typical day. This included: walking, moderate physical activity, vigorous physical activity, driving, computer use (excluding work-related use), and watching TV. Participants were requested to estimate an average time if the duration of activities differed considerably throughout the week. Variables relating to physical activity were recorded in minutes, while sedentary behaviours were recorded in hours. At the stage of data collection, answers indicating either negative values (<0) or duration exceeding 24 hours (1440 minutes) were rejected. By summing time spent on the three activities from each category, I additionally derived variables representing total daily physical activity and total daily sedentary behaviour duration. For both variables, I excluded cases where the sum exceeded 24 hours.

4.2.2.2 Predictors of interest in study B

I selected six variables reflecting both objective and subjective aspects of social engagement: frequency of family and/or friend visits (made and received), satisfaction with family relationships, satisfaction with friendships, frequency of opportunities to confide in someone, loneliness, and participation in social activities. I grouped responses relating to frequency of interactions into four categories: never, once every few months to once a month, one to four times a week, and daily or almost daily. I dichotomised satisfaction with relationships to “satisfied” and “not satisfied”. Similarly, I dichotomised experience of loneliness into “lonely” and “not lonely”. Regarding participation in social activities, I distinguished seven categories: reporting no engagement in social activities, attending a sports club or gym, going to a pub or a social club, participation in a religious group, attending an adult education class, other group activity, and engagement in multiple group activities.

4.2.2.3 Measures of cognitive performance

I chose four baseline cognitive tasks as outcome measures for both studies A and B. This included: reaction time, verbal-numerical reasoning (referred to as "Fluid Intelligence" in the UK Biobank), visual memory (referred to as "Pairs Matching") and prospective memory. Reasoning and prospective memory tasks had been added to the assessment procedure at a later stage of recruitment, and therefore were completed by fewer participants (254).

The reaction time task (data field 20023) included 12 rounds (4 training rounds, 8 trials) of card-matching, based on the game "Snap". Participants were presented with two cards at a time and asked to press a button as quickly as possible if they were identical. Performance on the task was measured as the average response time across eight trial rounds in milliseconds. Times under 50 milliseconds and over 2000 milliseconds were considered invalid, and thus excluded from the dataset.

The verbal-numerical reasoning task (data field 20016) involved answering 13 multiple choice logic/reasoning-type questions, within a 2-minute time limit. Performance was measured as the unweighted sum of correctly answered questions, with a maximum of 13 points.

In the visual memory task (data field 399) participants were presented with a set of matching pairs of cards and requested to memorise their positions. The cards were then turned over, and the subjects asked to select matching pairs in as few attempts as possible. The task included rounds with three and six pairs of cards, with performance measured as the number of errors in each round. In order to avoid a ceiling effect, for my analyses I only used results from the six-pair round, as it was more likely for participants to make an error.

For the prospective memory task (data field 20018), an initial instruction was given early in the cognitive assessment section. Participants were informed that at a later stage they will be shown four coloured shapes and asked to touch a blue square. Instead, however, they are to touch an orange circle. Originally, performance was grouped into three categories: incorrect response/task skipped, correct on first attempt, correct on second attempt. However, for the

purpose of my analyses, I dichotomised the outcome according to whether participants correctly responded on their first attempt or not.

All cognitive tasks were completed using a touchscreen. Additional information on cognitive testing in the UK Biobank can be found in the online Data Showcase under the "Cognitive function" category (category ID: 100026). Previous publications have described cognitive data from the repository with more detail, including information on test reliability and validity (255-258).

4.2.2.4 Covariates

Based on previous research, I identified factors simultaneously associated with physical activity, social engagement and cognition, which could act as confounders (i.e. lead to observing spurious associations between factors of interest and the outcomes). Firstly, I considered demographics: age (as a continuous variable, in years), sex, educational attainment, and the Townsend deprivation index score. I dichotomised education based on whether participants reported attainment of a college or university degree. The Townsend deprivation index is a measure of material deprivation based on rates of unemployment, car and home ownership, and household overcrowding in a given population (259). Each participant was assigned a deprivation index score at recruitment, based on a preceding national census. Negative values indicate relative affluence, while positive values indicate material deprivation.

Secondly, I included factors related to general health status and functioning: self-reported walking pace (three categories: brisk, steady/average, or slow), activity-limiting disability (dichotomised into present or absent), subjective health rating (four categories: excellent, good, fair, or poor), and BMI (entered into the analysis as a continuous variable). Disability was determined based on responses to a question on employment status (data field 6142). In answer to this questionnaire item, participants were able to select multiple response options, one of which was "unable to work because of sickness or disability". I considered an activity-limiting disability to be present if this response was selected either on its own or in conjunction with another option, for example, "retired" or "unemployed".

I additionally considered the presence of specific conditions that have been associated with vascular dementia and poorer cognitive outcomes in stroke populations (260-262): hypertension, diabetes and atrial fibrillation. In relation to mental health, I assumed a relevant association of physical activity patterns, social engagement and cognition with depression (263-265). I identified participants with a history of depressive episodes applying a method used in a previous UK Biobank-based study, combining responses that jointly indicated experiencing a period of feeling down, depressed, disinterested or unenthusiastic for at least two weeks, and seeking professional help (266).

The third category of covariates included lifestyle factors: frequency of alcohol intake (never/special occasions only, one to three times a month, one to four times a week, and daily/almost daily), and smoking (never, previous or current). These were treated as ordinal variables. Finally, I included two factors relating to the index cerebrovascular event: type (stroke or TIA), and time elapsed between the most recent stroke/TIA and baseline assessment, as a continuous variable measured in years.

4.2.3 Statistical analysis

4.2.3.1 Procedures

Following data inspection, I noted a positive skew for the following variables: walking, moderate physical activity, vigorous physical activity, driving, computer use, and total daily physical activity (six variables of interest in Study B); as well as for reaction time and visual memory data (outcomes in both studies). I used a natural log transformation to correct for this. For all variables apart from reaction time, I preceded the transformations by adding a constant (one) to all task results, to accommodate for possible values of zero. I performed the data transformations using the International Business Machines Corporation (IBM) Statistical Package for the Social Sciences (SPSS), version 24.

In both studies, I conducted a series of regression models to investigate associations between predictors of interest and performance on cognitive tasks. I used linear regression for three outcomes – reaction time, verbal-numerical reasoning and visual memory, and logistic regression for prospective memory.

The types of models that I developed are described below, separately for Studies A and B. I performed all analyses in Stata, version 14.2 (StataCorp LLC).

Due to a high proportion of missing data for some of the considered predictor and outcome measures (even reaching 2/3), I determined that data imputation would not be appropriate (267), and applied a complete-case analysis approach instead. Therefore, sample sizes varied by model and cognitive task. In studies based on large samples, such as included in the UK Biobank, statistically significant results can be obtained for even trivial effects (268). To account for this, as well as for multiple testing, I adopted a relatively strict approach, adjusting the traditionally recognised significance threshold according to the number of predictors in the most complex model (20 variables). Specifically, I set the threshold to 0.003 (0.05/20).

4.2.3.2 Study A models

I developed two groups of models in Study A. The first group accounted for the types of physical activity and sedentary behaviour that were reported. There were three models where the three types of physical activity were entered together, and three models where the three types of sedentary behaviour were entered together. In both cases, the first of these three models was unadjusted, incorporating only predictors of interest. The second was a partially adjusted model, where I included demographics (age, sex, education, and deprivation). The third model was fully adjusted, where I additionally accounted for health-, lifestyle- and stroke-related factors (walking pace, disability, subjective health rating, BMI, hypertension, diabetes, atrial fibrillation, depression, alcohol intake, smoking status, type of index cerebrovascular event, and time elapsed between the event and baseline assessment). Finally, I developed a fourth type of model, labelled “complete”, combining all types of physical activity and sedentary behaviour, and all potential covariates.

In the second group, I used total daily physical activity and total daily sedentary duration as predictors of interest. Here, I also developed four types of models, equivalent to the ones described above for the first group. In both model groups, I conducted separate unadjusted, partially adjusted, fully adjusted and complete models for each of the four cognitive tasks.

4.2.3.3 Study B models

The development of models in Study B reflected that of Study A. I created separate unadjusted, partially adjusted and fully adjusted models for each individual proxy of social engagement. I then incorporated all proxies into a single, complete model, adjusting for all considered covariates. I conducted every type of model for each of the cognitive tasks.

4.2.3.4 Presentation of results

I aimed to present my results in a manner that is both informative and intuitive. For analyses based on linear regression, I reported unstandardised coefficients (betas) to clearly indicate changes in values of the outcome measures. For logistic regression analyses, I reported odds ratios (OR).

I recognised that interpretation of coefficients is not straightforward where studied variables were log-transformed. In such cases, I provided examples to represent the strength of the reported associations. Underlying calculations were based on: i) percent change in the outcome for every one-unit change in the predictor, where only the outcome was log-transformed, ii) unit change in the outcome for every 1% change in the predictor, where only the predictor was log-transformed, iii) percent change in the outcome for every 1% change in the predictor, where both the predictor and outcome were log-transformed (269). For complete models, I additionally presented the results graphically. However here, for linear regression analyses, I used standardised coefficients to allow visual comparisons across different predictors of interest.

4.3 Results and discussion

I identified 8391 participants with a self-reported history of stroke or TIA. Table 4-1 presents descriptive statistics for variables relevant to both studies: baseline characteristics of the participant sample, incorporated in conducted analyses as covariates, and measures of cognitive performance, constituting study outcomes.

Table 4-1 Baseline characteristics of study sample and descriptive statistics for performance on cognitive tasks.

Variables	
Demographics	
Age, years	
Mean (SD)	61.1 (6.6)
Sex	
Female	3508/8391 (41.8%)
Degree-level education	1772/8190 (21.6%)
Missing data	201
Townsend deprivation score (higher: more deprived)	
Mean (SD)	−0.5 (3.5)
Missing data	10
Health-related factors	
Walking pace	
Brisk	1724/8104 (21.3%)
Steady/average	4002/8104 (49.4%)
Slow	2378/8104 (29.3%)
Missing data	287
Disability	1488/8347 (17.8%)
Missing data	44
Subjective health rating	
Excellent	339/8288 (4.1%)
Good	3192/8288 (38.5%)
Fair	3171/8288 (38.3%)
Poor	1586/8288 (19.1%)
Missing data	103
BMI	
Mean (SD)	28.9 (5.2)
Missing data	129
Comorbidities	
Hypertension	4811/8391 (57.3%)
Diabetes	1187/8391 (14.2%)
Atrial fibrillation	253/8391 (3.0%)
History of depressive episodes	644/5566 (11.6%)
Missing data	2825

Table 4-1 Baseline characteristics of study sample and descriptive statistics for performance on cognitive tasks. *Continued*

Variables	
Lifestyle factors	
Alcohol intake frequency	
Never/special occasions only	2473/8364 (29.6%)
Monthly	900/8364 (10.7%)
Weekly	3384/8364 (40.5%)
Daily/almost daily	1607/8364 (19.2%)
Missing data	27
Smoking status	
Never	3502/8316 (42.1%)
Previous	3533/8316 (42.5%)
Current	1.281/8316 (15.4%)
Missing data	75
Stroke-related factors	
Type of cerebrovascular event	
Stroke	6773/8391 (80.7%)
TIA	1618/8391 (19.3%)
Time from stroke/TIA to baseline assessment, years	
Mean (SD)	7.3 (7.0)
Missing data	622
Cognitive task performance	
Reaction time, milliseconds	
Mean (SD)	613.5 (151.7)
Missing data	193
Verbal-numerical reasoning, points (range 0 to 13)	
Mean (SD)	5.3 (2.1)
Missing data	5761
Visual memory, errors	
Mean (SD)	4.6 (3.7)
Missing data	74
Prospective memory	
Correct response on first attempt	1855/2839 (65.3%)
Missing data	5552

BMI indicates body mass index; SD, standard deviation; TIA, transient ischaemic attack.

4.3.1 Study A results

Descriptive statistics for predictors of interest in Study A, relating to self-reported duration of physical activities and sedentary behaviours on a typical day, are presented in Table 4-2. Across all conducted models, sample sizes ranged from 1571 to 7927 participants. Given multiple analyses, producing a high volume of results, I have focused on describing only key findings. Associations for all variables of interest are included in Appendix 4, Supplemental Tables 1 to 8.

4.3.1.1 Predictors of reaction time in Study A

In unadjusted models, faster reaction times were associated with longer reported daily duration of: walking (beta = −0.011, 99.7% CI: −0.019 to −0.003), vigorous physical activity (beta = −0.006, 99.7% CI: −0.012 to −0.001), driving (beta = −0.032, 99.7% CI: −0.047 to −0.017), and computer use (beta = −0.035, 99.7% CI: −0.048 to −0.023), as well as total physical activity (beta = −0.012, 99.7% CI: −0.018 to −0.005). Time spent watching TV was the only factor associated with slower reaction times (beta = 0.008, 99.7% CI: 0.005 to 0.012).

After adjusting for demographics, vigorous physical activity was no longer a predictor of reaction time, while associations remained similar for: walking (beta = −0.011, 99.7% CI: −0.019 to −0.003), driving (beta = −0.020, 99.7% CI: −0.035 to −0.005), computer use (beta = −0.028, 99.7% CI: −0.040 to −0.015), watching TV (beta = 0.006, 99.7% CI: 0.002 to 0.009), and total physical activity duration (beta = −0.011, 99.7% CI: −0.018 to −0.005).

In fully adjusted models, additionally accounting for health-, lifestyle- and stroke-related factors, computer use was the only significant predictor of task performance (beta = −0.023, 99.7% CI: −0.039 to −0.008). To illustrate the magnitude of this association, with every 50.0% increase in computer use duration, reaction time decreased by 0.9%. In the complete model (Figure 4-1), simultaneously including all types of activities/behaviours and covariates, the results were similar – only computer use showed a weak, significant association with reaction time (beta = −0.024, 99.7% CI: −0.042 to −0.006). Across all models, I found no associations between task performance and either moderate physical activity or total sedentary time (Figure 4-5 A presents complete model).

Table 4-2 Descriptive statistics for self-reported duration of physical activities and sedentary behaviours, performed during a typical day.

Variables	
Physical activities (minutes on a typical day)	
Walking	
Mean (SD)	54.7 (72.9)
Median (IQR)	30.0 (40.0)
Missing data	1469
Moderate activity	
Mean (SD)	54.4 (76.2)
Median (IQR)	30.0 (50.0)
Missing data	1540
Vigorous activity	
Mean (SD)	19.5 (39.0)
Median (IQR)	0.0 (30.0)
Missing data	1232
Total active time (minutes on a typical day)	
Mean (SD)	130.5 (142.3)
Median (IQR)	90.0 (125)
Missing data	2545
Sedentary behaviours (hours on a typical day)	
Driving	
Mean (SD)	0.7 (1.1)
Median (IQR)	0.0 (1.0)
Missing data	238
Computer use	
Mean (SD)	1.0 (1.5)
Median (IQR)	0.0 (2.0)
Missing data	205
Watching TV	
Mean (SD)	3.6 (2.2)
Median (IQR)	3.0 (3.0)
Missing data	164
Total sedentary time (hours on typical day)	
Mean (SD)	5.2 (2.8)
Median (IQR)	5 (4)
Missing data	465

IQR indicates interquartile range; SD, standard deviation.

4.3.1.2 Predictors of verbal-numerical reasoning in Study A

In unadjusted models, better verbal-numerical reasoning task performance was associated with longer daily computer use (beta = 0.546, 99.7% CI: 0.342 to 0.749), while worse performance was associated with watching TV (beta = −0.191, 99.7% CI: −0.248 to −0.133) and total duration of sedentary behaviour (beta = −0.068, 99.7% CI: −0.113 to −0.022). Although reduced in magnitude, these associations remained significant after adjusting for demographics: beta = 0.396, 99.7% CI: 0.192 to 0.599, for computer use; beta = −0.133, 99.7% CI: −0.191 to −0.075, for watching TV; and beta = −0.048, 99.7% CI: −0.092 to −0.003, for total sedentary time.

In fully adjusted models, computer use and watching TV were the only behaviours that predicted task performance. On average, for every 50% increase in computer use duration, the task score increased by 0.1 of a point (beta = 0.336, 99.7% CI: 0.108 to 0.564), while the score decreased by a similar amount with every 1-hour increase in watching TV (beta = −0.100, 99.7% CI: −0.169 to −0.032). Associations were comparable in the complete model (Figure 4-2): beta = 0.275, 99.7% CI: 0.018 to 0.533, for computer use; and beta = −0.116, 99.7% CI: −0.194 to −0.038, for watching TV.

Regarding total sedentary behaviour duration, after adjusting for all covariates, the initial association with verbal-numerical reasoning task scores was no longer significant. However, in the complete model (Figure 4-5 B), additionally including total physical activity duration among predictors, the association approached the adopted significance threshold (beta = −0.056, 99.7% CI: −0.115 to 0.003, p = 0.005), suggesting that with an increase in time spent sedentary throughout an average day, task performance may have marginally worsened. Across all conducted models, physical activity did not predict task scores – neither when considering particular types of activities nor total daily active time.

4.3.1.3 Predictors of visual memory in Study A

In unadjusted models, I found associations between the number of errors in the visual memory task and moderate physical activity duration and total physical activity duration, nearing the significance threshold (p = 0.003 and p = 0.004, respectively). In both cases, longer activity duration was weakly associated with

more errors: beta = 0.019, 99.7% CI: 0.000 to 0.037, for moderate physical activity; and beta = 0.020, 99.7% CI: −0.001 to 0.040, for total activity.

After adjusting for demographics, my only finding was a trend for an association between task performance and watching TV, where more TV time predicted less errors (beta = −0.011, 99.7% CI: −0.022 to <0.001, p = 0.004). In the fully adjusted and complete models (Figure 4-3 and Figure 4-5 C), there were no significant associations between predictors of interest and visual memory task performance.

4.3.1.4 Predictors of prospective memory in Study A

In unadjusted models, I observed that a correct response on first attempt in the prospective memory task was more likely with longer daily computer use (OR = 1.316, 99.7% CI: 1.066 to 1.625), and less likely with more time spent watching TV (OR = 0.908, 99.7% CI: 0.857 to 0.962). Associations with task performance remained similar after adjusting for demographics: OR = 1.251, 99.7% CI: 1.005 to 1.557, for computer use; and OR = 0.939, 99.7% CI: 0.884 to 0.996, for watching TV. However, none of the sedentary behaviours, nor their total duration, predicted the likelihood of a correct response in the fully adjusted and complete models (Figure 4-4 and Figure 4-5 D). Physical activity, considered both in terms of distinct types and total duration, was not associated with task performance in any of the models I conducted.

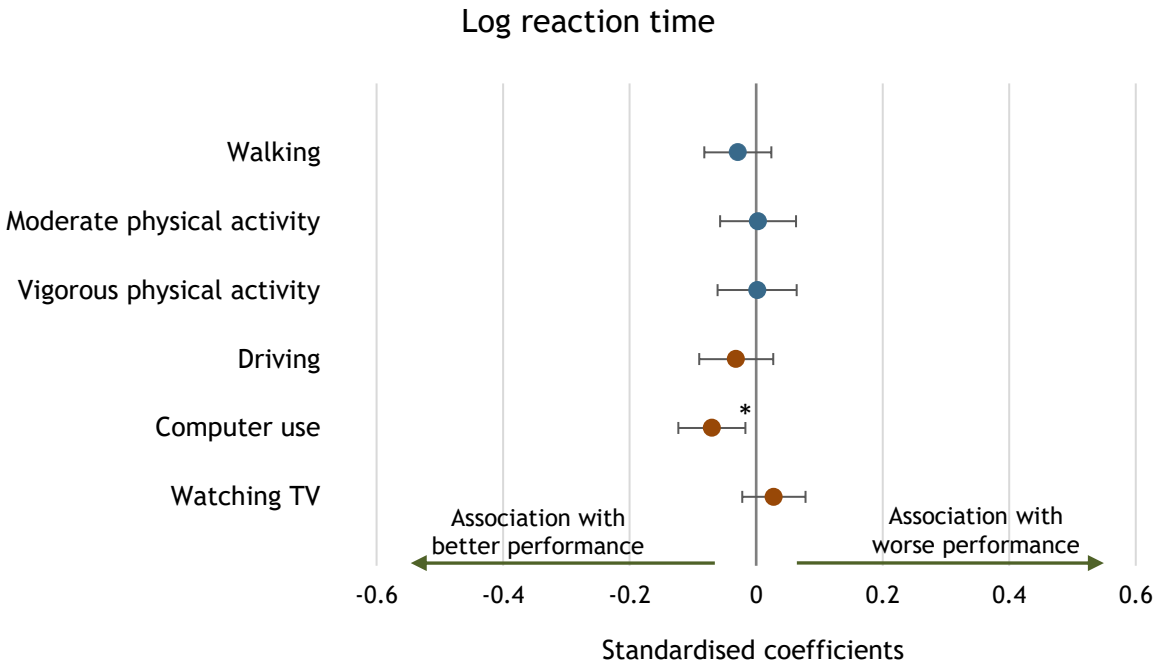


Figure 4-1 Associations of log reaction time with daily duration of physical activity and sedentary behaviour according to type in a complete model, with 99.7% CI.

Notes: Apart from watching TV, reported duration for predictors was log-transformed;
*p < 0.003.

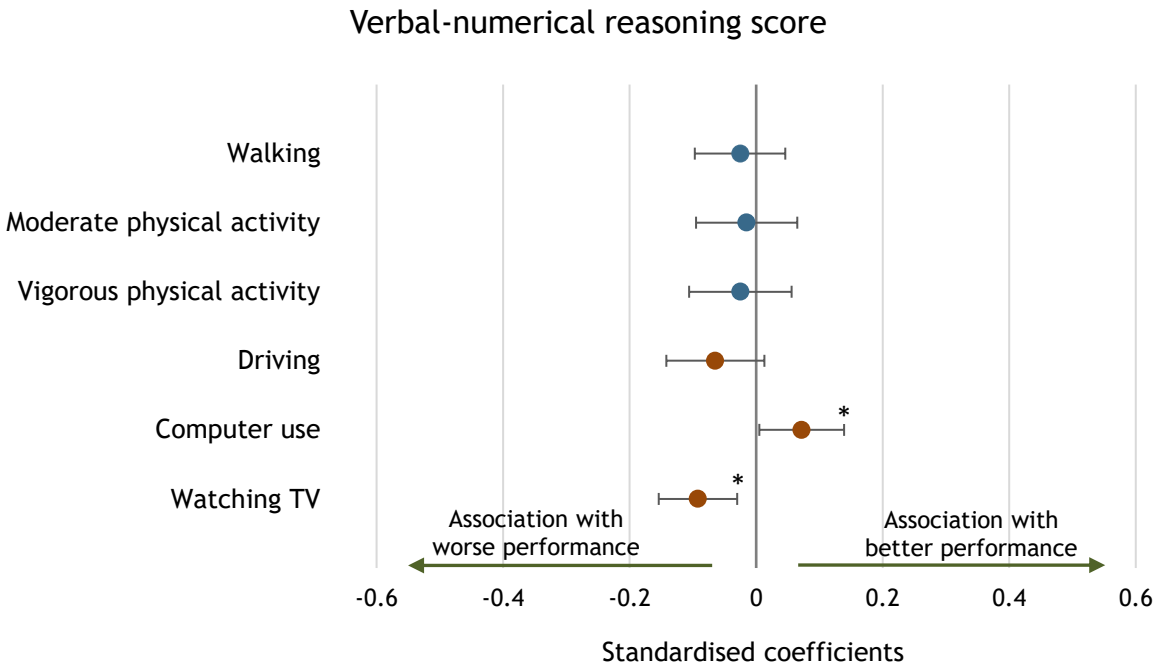


Figure 4-2 Associations of verbal-numerical reasoning task scores with daily duration of physical activity and sedentary behaviour according to type in a complete model, with 99.7% CI.

Notes: Apart from watching TV, reported duration for predictors was log-transformed;
*p < 0.003.

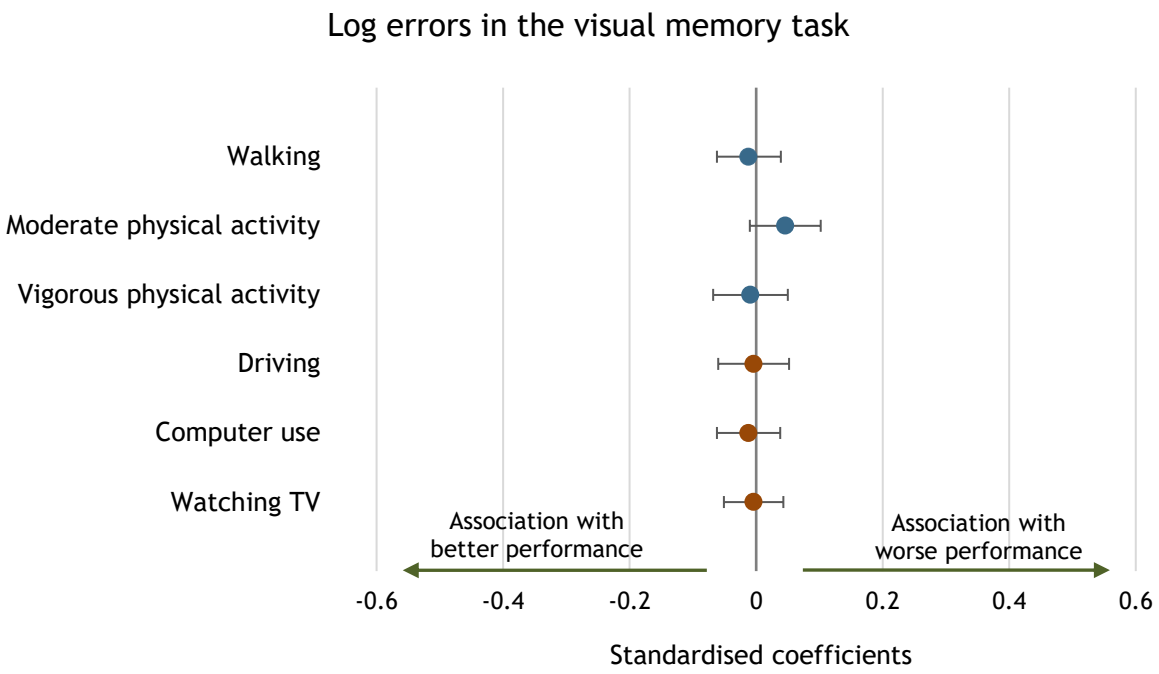


Figure 4-3 Associations of log errors in the visual memory task scores with daily duration of physical activity and sedentary behaviour according to type in a complete model, with 99.7% CI.

Notes: Apart from watching TV, reported duration for predictors was log-transformed.

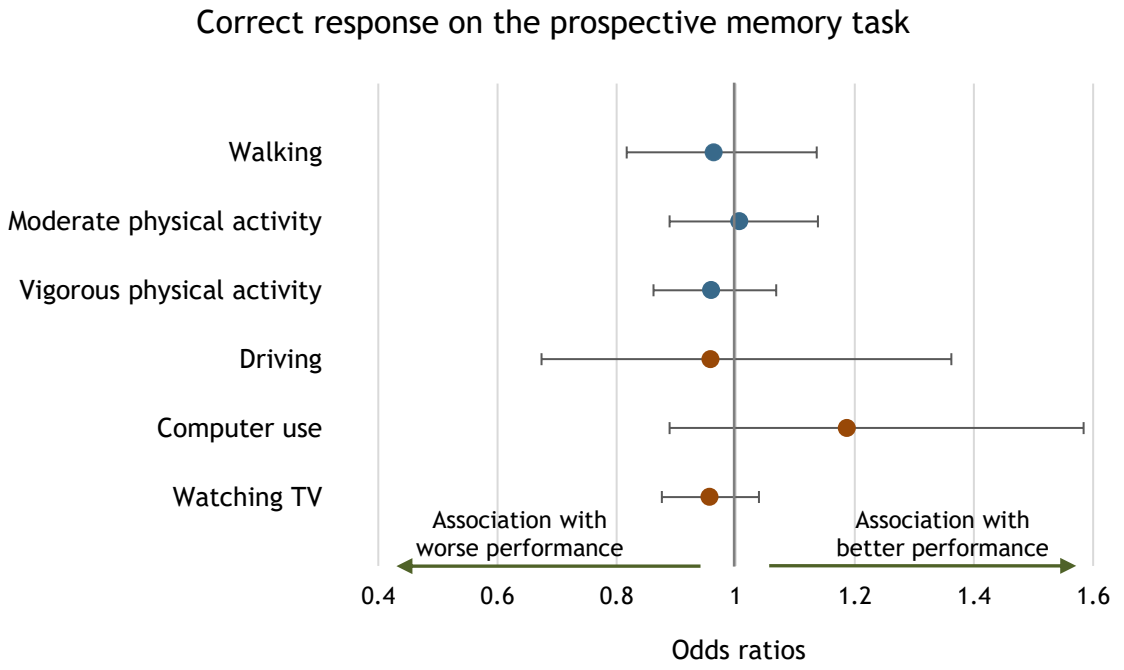


Figure 4-4 Associations of a correct response on the prospective memory task with daily duration of physical activity and sedentary behaviour according to type in a complete model, with 99.7% CI.

Notes: Apart from watching TV, reported duration for predictors was log-transformed.

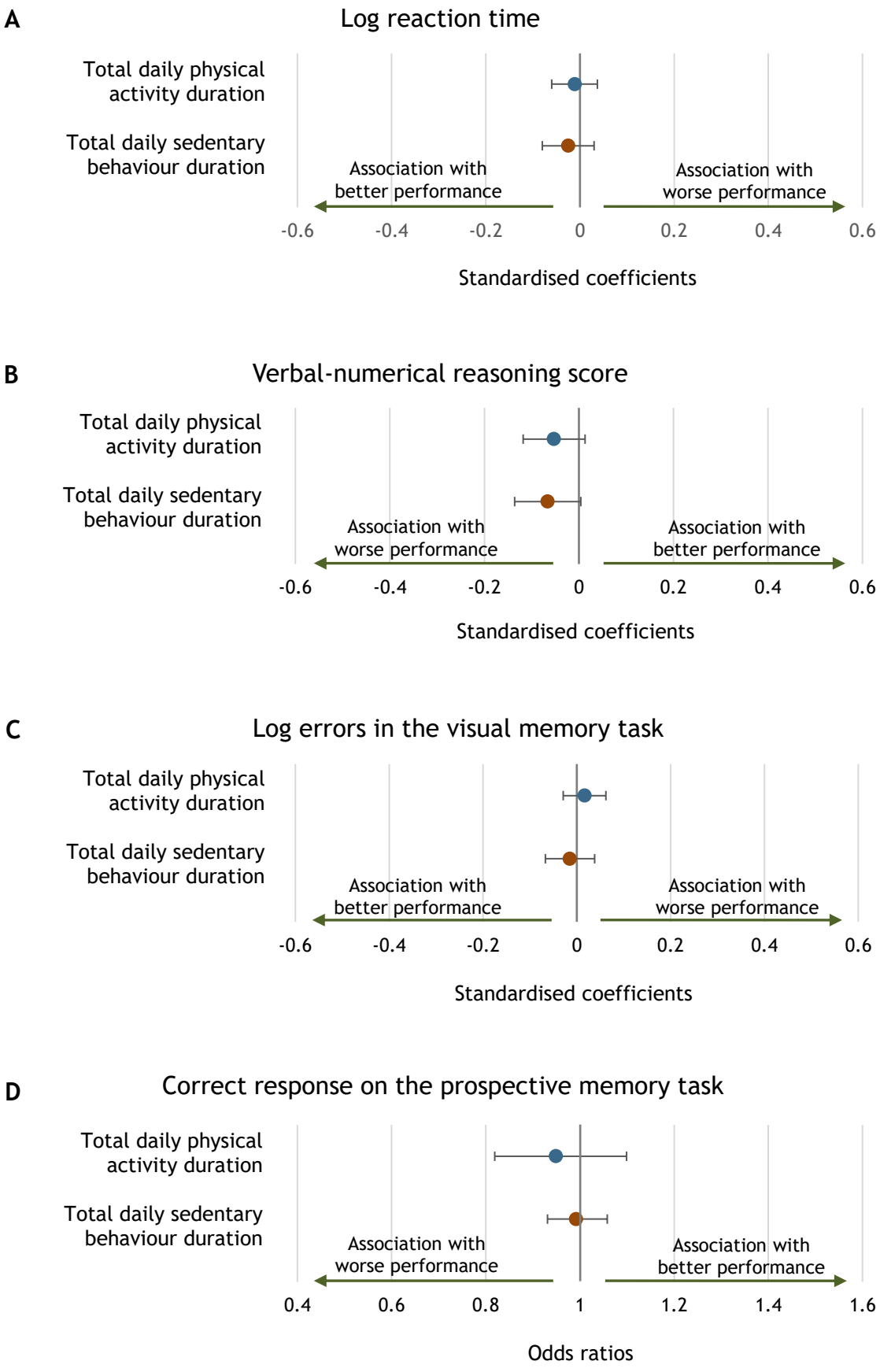


Figure 4-5 Associations of log reaction time (A), verbal-numerical reasoning scores (B), log errors in the visual memory task (C), and a correct response on the prospective memory task (D) with log of total daily physical activity duration and total daily sedentary behaviour duration in complete models, with 99.7% CI.

4.3.2 Study A discussion

In this cross-sectional study in a sample of stroke survivors from the UK Biobank, I found few significant task-specific associations between cognitive performance and measures of physical activity and sedentary behaviour, all of which were very weak. The most consistently demonstrated effects were for computer use and watching TV. Some of the findings seem to oppose expectations arising from previous research, including a similar study conducted in the whole UK Biobank population (270).

In that study, with a main focus on sedentary behaviour, the authors reported an inverse association between driving time and cognitive performance across all considered tasks, which my results did not reflect. A plausible explanation relates to the characteristics of my participant sample. Following stroke, many people discontinue driving due to motor, perceptual and/or cognitive deficits, or even low confidence (271-273). As follows, participants who reported regular driving in my study may also be individuals who on average experienced less severe post-stroke impairments, including less affected cognition.

Regarding the remaining two types of sedentary behaviour, conclusions from our studies generally seem to align. I found that watching TV was modestly associated with worse performance on reaction time, prospective memory and verbal-numerical reasoning tasks. Yet, only in case of the latter the effect remained significant after adjusting for all covariates. There are two potential explanation for this: i) in part, sedentary behaviour detrimentally affects cognition through impairing health overall, ii) people in poor health/feeling unwell tend to spend more time sedentary. Thus, once health-related factors are controlled for, some association may no longer be apparent.

Research on sedentary behaviour and its effects has long focused on watching TV as a central activity of interest. Many studies have demonstrated the negative impact of extensive TV viewing on health outcomes, with evidence also expanding for a deleterious association with cognition and mood (274-277). These findings may have encouraged the supposition that all forms of sedentary behaviour are uniformly detrimental across outcomes. However, this may not be necessarily the case, as demonstrated for the unselected UK Biobank population

and here, where computer use was associated with better cognitive performance in most conducted models (none relating to visual memory).

My results are in line with conclusions from a different area of research, focusing on behaviours that help preserve cognitive function in older age. Multiple reports suggest a protective effect of engaging in mentally stimulating activities, many of which are performed while being sedentary (e.g. playing cards, doing puzzles, reading books, sewing) (278-280). Moreover, in research on sedentary behaviour, a recent overview of systematic reviews indicated that a favourable association of computer and internet use with cognitive function in older adults was one of the key and novel findings (205).

However, in the current state of knowledge, it is important to interpret these observations with caution. There is some evidence from randomised controlled trials in healthy older adults to support a causal relationship between computer training and improved cognitive performance; yet, in a meta-analysis the overall effect size was small, while associations were found to differ depending on the form of intervention (281). Namely, unsupervised home-based training was concluded to be ineffective.

It therefore seems that at least in part the positive association between computer use and cognition could be attributed to individuals with a higher level of cognitive function being more likely to engage in this activity. In the case of my study, it is also possible that participants who used a computer more frequently had greater ease in completing cognitive tasks on a touchscreen – an advantage that would not have been present if the assessment were paper-based.

In view of the above arguments and discussed evidence, it is perhaps the findings on physical activity that are more surprising here, with hardly any support for the well-documented positive effect on cognitive performance. The only such observation related to faster reaction times being associated with longer duration of daily walking, vigorous activity and total physical activity. These effects seem consistent with conclusions from the previously described meta-analysis, where participation in structured physical activity training increased processing speed (233).

However, in my study, none of the associations remained significant after adjusting for all covariates. It is therefore plausible that the effects of physical activity were driven by more active participants being in better health, which in turn translated to faster processing speed. This interpretation would not yet account for the null results across other cognitive tasks, or the even more unexpected finding that moderate and total daily physical activity were weakly associated with more errors in the unadjusted model for visual memory.

One explanation that aligns with these collective observations relates to occupation, particularly in light of physical activity being reported for a whole, typical day, and not just regarding leisure time. Previous studies have found that in contrast to jobs characterised by high intellectual demands, and human interaction and communication, occupations based on physical activity are associated with poorer cognitive function (282-285). This effect appears to be independent of education and deprivation, and so sociodemographic factors included in my adjusted models would not have controlled for it. As follows, it is possible that the opposing effects of occupational and leisure time physical activity jointly manifested as a neutral association with cognitive performance.

Another important consideration relates to use of self-reported measures. Being prone to recall and social desirability bias, it is difficult to ascertain how closely they reflected actual energy expenditure, particularly as studies comparing subjectively and objectively assessed physical activity indicate correlations differing not only in strength, but also in direction (286).

Lastly, the association between physical activity and poorer visual memory performance brings attention to the fact that all findings for this task diverged from the overall pattern of results. This may be due to the test's specific properties. Studies describing the UK Biobank cognitive assessment found that the visual memory task had: poor test-retest reliability; compared to other tasks, had one of the weakest associations with a general cognitive ability score; and after adjusting for age, only a weak association with a validated reference test (257, 258).

4.3.2.1 Clinical implications

As noted above, although I observed relatively consistent associations with cognitive performance for computer use and watching TV, these effects were very weak. Thus, if considered in isolation, my results are insufficient to conclude that the average daily duration of these behaviours is linked to clinically meaningful differences in cognitive function. However, cumulatively interpreting the available evidence, it seems that reducing mentally passive sedentary behaviour by replacing it with mentally active behaviour and some form of physical activity is plausibly of therapeutic value and can be recommended (279, 287, 288).

In relation to breaking-up sedentary time, it seems that at least some gains can be expected in terms of physical health and mood (205, 289), which are in turn associated with cognitive function. As a group, stroke survivors also seem to be among the more likely to benefit from such a change to habits, as the positive effects of interrupting prolonged periods of sitting with light-intensity physical activity are found to be more pronounced in physically inactive individuals (236).

Moreover, as a potential intervention, introducing mentally and physically active behaviours to break up mentally passive sedentary time is worth considering in view of its practical advantages. Strategies based on this approach could be easily tailored to suit individual likes and interests, functional status, and living conditions, and subsequently be implemented by stroke survivors without the need for professional supervision, purchasing specialist equipment or travel.

Many of the presented arguments in favour of breaking-up sedentary behaviour have also been advocated by investigators involved in the RECREATE study (REduCing sedentaRy bEhaviour After sTrokE) (290). Insights gained from this project, beginning from systematic reviews and ending on a multicentre cluster randomised controlled trial, may indeed find a place in future stroke-specific clinical guidelines. Presently, however, there is no indication that assessed outcomes will encompass cognitive function (291), and so it seems that the need for studies addressing the role of sedentary behaviour in shaping post-stroke cognitive outcomes has not yet been met.

4.3.2.2 Research implications

Results of the analyses partially supported my initial assumptions, which may have a direct application to future research. Firstly, it seems that sedentary behaviour is associated with cognitive performance independently of physical activity. As follows, both categories of activity patterns should be accounted for in prediction model studies, and both can be considered as potential targets for intervention.

Secondly, distinguishing between different types of sedentary behaviour is a valid approach, particularly regarding whether their nature is mentally passive or active, as this could entail opposing effects on cognition. However, implementing this notion into research practice can present a considerable challenge, as behaviours plausibly exist on a continuum, and accurately classifying some of them will require a high level of detail. For example, social media can be accessed to participate in discussion forums or to watch short videos for entertainment purposes. Arguably, the former is more cognitively stimulating than the latter, yet both behaviours would fall under the broad category of mobile device use.

Fortunately, in the same review where the distinction between the two types of sedentary behaviour was first proposed, an assessment framework was also presented (193). The authors suggested considering sedentary behaviour across three different contexts – occupation, leisure and transport – and provided mentally active and passive examples relevant to each. Conceptually, this novel approach seems highly useful in the context of identifying modifiable predictors of cognitive function, although its feasibility, validity and reliability are yet to be determined in future studies.

4.3.2.3 Limitations

One of the main limitations of my study was the aforementioned use of self-reported measures for key predictors of interest, which introduces concerns around estimate accuracy. Ideally, physical activity would be assessed using an accelerometer, however, for the UK Biobank such data was not collected until seven years after recruitment had begun and involved only a fifth of the sample (292). A similar issue relates to identifying eligible participants for my study.

Compared to an objective source (e.g. hospital records), relying on self-reported history of stroke or TIA may increase the risk of both missing relevant cases, as well as including false positives. In a systematic review of the accuracy of participant self-report of stroke, the latter problem was indicated as being of particular concern for large prospective studies, such as the UK Biobank, where healthy individuals are overrepresented, and stroke prevalence is likely to be low (293). At the same time, however, based on national stroke prevalence data and allowing for the "healthy cohort effect", the authors of the review estimated that the true prevalence of stroke in the UK Biobank is plausibly under 2%. This aligns with the 1.4% stroke prevalence recorded for the cohort based on self-report.

Due to sample characteristics, there are also potential concerns around generalisability, as with many UK Biobank-based studies. Participants in my study were on average younger, had a higher education and a lower comorbidity burden than an unselected stroke population (for comparison see 294). Yet, both inactivity and cognitive problems are likely to be more prevalent in a 'real world' group of stroke survivors, and the associations observed here may be exacerbated in a dedicated stroke cohort.

A further, related limitation is that I did not have access to data regarding stroke-related factors, recognised as predictors of cognitive outcome, including: stroke type, acute symptoms, acute physiology and, most importantly, stroke severity (23). I attempted to partially adjust for the latter by including disability among model covariates, serving as a proxy for post-stroke functional deficits (295). In hindsight, as discussed previously, I also acknowledge the relevance of controlling for occupation type. However, given the coding of this variable, an intention to include it in the analyses would have introduced a high level of additional complexity to my study.

In view of characterising the relationship between activity patterns and post-stroke cognitive function, it is moreover important to note that I had not comprehensively addressed the issue of time elapsed between index stroke or TIA and baseline assessment. While I controlled for this factor in the fully adjusted and complete models, I did not determine whether it moderated the associations between physical activity, sedentary behaviour and cognitive

outcomes. Identifying in what period following stroke activity patterns are particularly relevant to cognitive performance may significantly contribute to planning interventions based on modifying habitual behaviour.

Another, crucial limitation to understanding the nature of studied relationships stems from the cross-sectional design of my study, precluding causal inference. Although repeat assessments for the UK Biobank were conducted between 2012 and 2013 (approximately 20300 participants), data on variables of interest were available for only a small percentage of my study sample (e.g. 290 stroke survivors completed the verbal-numerical reasoning task at follow-up). As these attrition rates may have introduced considerable bias, it did not seem feasible to conduct a longitudinal analysis.

Even with the adopted approach, the extent of missing data was a pronounced issue in my study, affecting even 2/3 of participants for some of the variables. This was, however, for the most part due to certain baseline assessment tasks and questions added at later stages of UK Biobank recruitment. Thus, I assumed that most of these data were likely missing completely at random, i.e. that there were no systematic differences between the missing values and the observed values (296).

Finally, use of the UK Biobank repository entailed relying on results of bespoke cognitive tasks as measures of cognitive performance, designed to allow brief assessment on a large scale, without the need for examiner supervision. This approach, although having practical advantages, poses a considerable challenge to comparing the developed tests with standard cognitive measures, routinely administered in research and clinical practice. It also introduces some uncertainty regarding what cognitive functions are engaged in task completion.

4.3.2.4 Strengths

Alongside inherent limitations, use of the UK Biobank resource presented important advantages. I had access to data from a relatively large sample of stroke survivors – several times larger than most dedicated stroke cohorts. Further, the wealth of available variables made it possible to control for many

key factors simultaneously relevant to physical activity, sedentary behaviour and cognitive function.

It was also the variety of collected information that allowed me to adopt a relatively novel approach to considering effects of activity patterns – accounting for both physical activity and sedentary behaviour, and distinguishing between their different types. To my knowledge, this is the first study to address these aspects of routine behaviour in relation to cognitive function among stroke survivors, who as a group are at greater risk of both inactivity and cognitive impairment.

4.3.2.5 Future directions

The relationship of post-stroke cognition with habitual physical activity and sedentary behaviour appears far from unravelled, considering: the scarcity of existing evidence on this specific topic; the limitations of the present study; and the complexity of associations, with multiple potential paths of influence, originating even prior to index stroke (232). As follows, there is a need for bespoke cohort studies in stroke populations, combining objective measures of physical activity and sedentary behaviour with self-report to classify types of the latter, and employing sensitive and widely recognised cognitive assessment tools. These studies will also need to account for a number of sociodemographic, health and lifestyle-related factors.

Regarding approaches to analysis, additional insights might be gained by accounting for mediation effects. Specifically, it could be determined how much of the overall effect of physical activity and sedentary behaviour on post-stroke cognition is explained by associations with physical health (e.g. insulin resistance or risk of recurrent stroke) and mood; and whether there is evidence for a more direct impact of activity patterns on neurological function.

A better understanding of the considered associations could in turn inform the design of future interventions. Here, it will be essential to maintain a balance between the potential to produce clinically meaningful cognitive gains and feasibility. The efficacy of any intervention will be of little value if in the long term it cannot be embedded in everyday lives of predominantly older

individuals, many of whom experience functional disability, comorbidity, fatigue and apathy.

Developing such studies as I suggest will undoubtedly require considerable resources, and consequently, funding constitutes a key issue. One route that could increase the likelihood of securing financial support could lead through broadening the scope of this research. As outlined at the beginning of the chapter, physical activity and sedentary behaviour affect health and well-being in many ways. If alongside cognition multiple other outcomes were to be assessed (e.g. mortality, non-fatal adverse events, functional independence, depression or fatigue), this could expand the range of organisations willing to fund projects on post-stroke physical activity and sedentary behaviour, particularly, as so far RECREATE appears as a single large-scale trial in this specific research landscape.

4.3.3 Study B results

The distribution among categories of the six considered social engagement proxies is presented in Table 4-3. Across all models I conducted for Study B, sample sizes ranged from 1873 to 8266 participants. Similarly as for Study A, in this chapter I summarised key findings from my analyses, while complete results are included in Appendix 5, Supplemental Tables 9 to 12.

4.3.3.1 Reaction time

In unadjusted models, faster reaction times were associated with monthly (beta = −0.052, 99.7% CI: −0.095 to −0.009) and weekly (beta = −0.046, 99.7% CI: −0.086 to −0.006) family/friend visits as compared to no visits, and participation in sports as compared to no reported social activities (beta = −0.031, 99.7% CI: −0.058 to −0.003). Slower reaction times were associated with loneliness (beta = 0.024, 99.7% CI: 0.008 to 0.041) and participation in religious group activity (beta = 0.055, 99.7% CI: 0.026 to 0.085).

After adjusting for demographics, associations with weekly family/friend visits (beta = −0.043, 99.7% CI: −0.083 to −0.003), loneliness (beta = 0.020, 99.7% CI: 0.003 to 0.037), and religious group activity remained significant (beta = 0.048, 99.7% CI: 0.019 to 0.077). In addition, faster reaction times were predicted by satisfaction with family relationships (beta = −0.051, 99.7% CI: −0.095 to −0.007).

In fully adjusted models, I found only two proxies of social engagement to be significant predictors of task performance. Monthly family/friend visits were associated with 6.5% faster reaction times (beta = −0.063, 99.7% CI: −0.119 to −0.007), while religious group activity – with 4.8% slower times (beta = 0.047, 99.7% CI: 0.011 to 0.084). However, in the complete model (Figure 4-6), I observed no significant associations between any proxies of social engagement and reaction time. Satisfaction with friendships and frequency of opportunities to confide in someone were the only variables that did not predict task performance in any of the models.

Table 4-3 Distribution among categories of social engagement proxies for study sample.

Variables	
Family/friend visits	
Never	284/8232 (3.4%)
Every few months/monthly	1441/8232 (17.5%)
Weekly	5241/8232 (63.7%)
Daily/almost daily	1266/8232 (15.4%)
Missing data	159
Family satisfaction	
Satisfied	2576/2818 (91.4%)
Missing data	5573
Friendship satisfaction	
Satisfied	2687/2804 (95.8%)
Missing data	5587
Loneliness	
Lonely	2060/8212 (25.1%)
Missing data	179
Opportunities to confide in someone	
Never	1638/8022 (20.4%)
Every few months/monthly	955/8022 (11.9%)
Weekly	1487/8022 (18.6%)
Daily/almost daily	3942/8022 (49.1%)
Missing data	369
Social activities	
None	2880/8333 (34.6%)
Sports club/gym	694/8333 (8.3%)
Pub/social club	1418/8333 (17.0%)
Religious group	602/8333 (7.2%)
Adult education class	151/8333 (1.8%)
Other group activities	809/8333 (9.7%)
Multiple group activities	1779/8333 (21.4%)
Missing data	58

4.3.3.2 Verbal-numerical reasoning

In unadjusted models, I found that better verbal-numerical reasoning task scores were associated with engagement in multiple social activities (beta = 0.530, 99.7% CI: 0.200 to 0.861), while worse scores were associated with reported loneliness (beta = −0.654, 99.7% CI: −0.938 to −0.370). In partially and fully adjusted models, loneliness remained as the only significant predictor of task performance, associated with an approximately 0.4-point decrease in scores (beta = −0.417, 99.7% CI: −0.697 to −0.136, and beta = −0.412, 99.7% CI: −0.750 to −0.074, respectively).

However, having combined all proxies of social engagement in the complete model (Figure 4-7), the association with loneliness was no longer significant, although close to the set threshold (beta = −0.345, 99.7% CI: −0.704 to 0.013, p = 0.004). Across all models, I observed no significant associations between verbal-numerical reasoning scores and family/friend visits, satisfaction with relationships, and frequency of opportunities to confide in someone.

4.3.3.3 Visual memory

Across all models, none of the studied proxies of social engagement significantly predicted the number of errors in the visual memory task. Figure 4-8 presents associations with variables of interest in the complete model.

4.3.3.4 Prospective memory

Across all models including individual proxies of social engagement, my results indicated that loneliness was the only significant predictor of performance on the prospective memory task. With reported loneliness, odds of a correct response on first attempt were lower by 30.5% (OR = 0.695, 99.7% CI: 0.532 to 0.907) in the unadjusted model, lower by 28.0% (OR = 0.720, 99.7% CI: 0.543 to 0.955) in the partially adjusted model, and lower by 33.9% (OR = 0.661, 99.7% CI: 0.461 to 0.947) in the fully adjusted model. However, in the complete model (Figure 4-9), the association with loneliness was no longer significant.

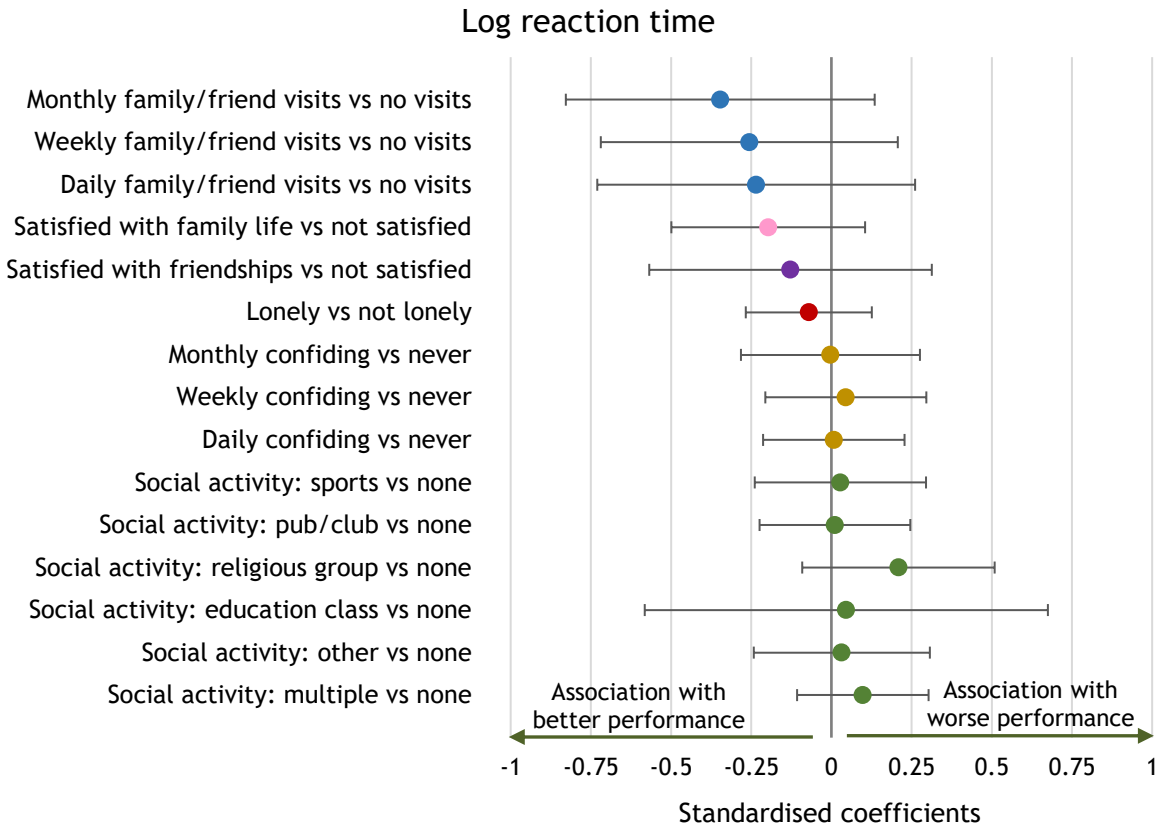


Figure 4-6 Associations of log reaction time with proxies of social engagement in a complete model, with 99.7% CI.

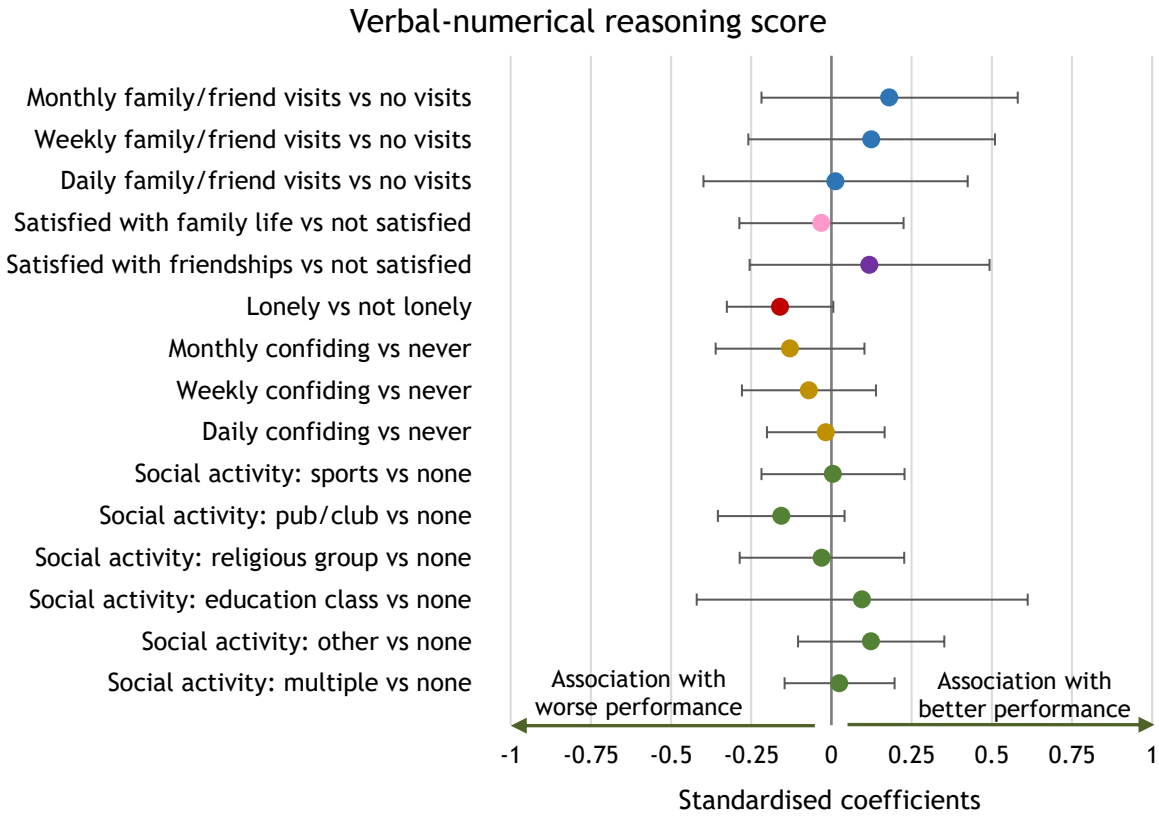


Figure 4-7 Associations of verbal-numerical reasoning task scores with proxies of social engagement in a complete model, with 99.7% CI.

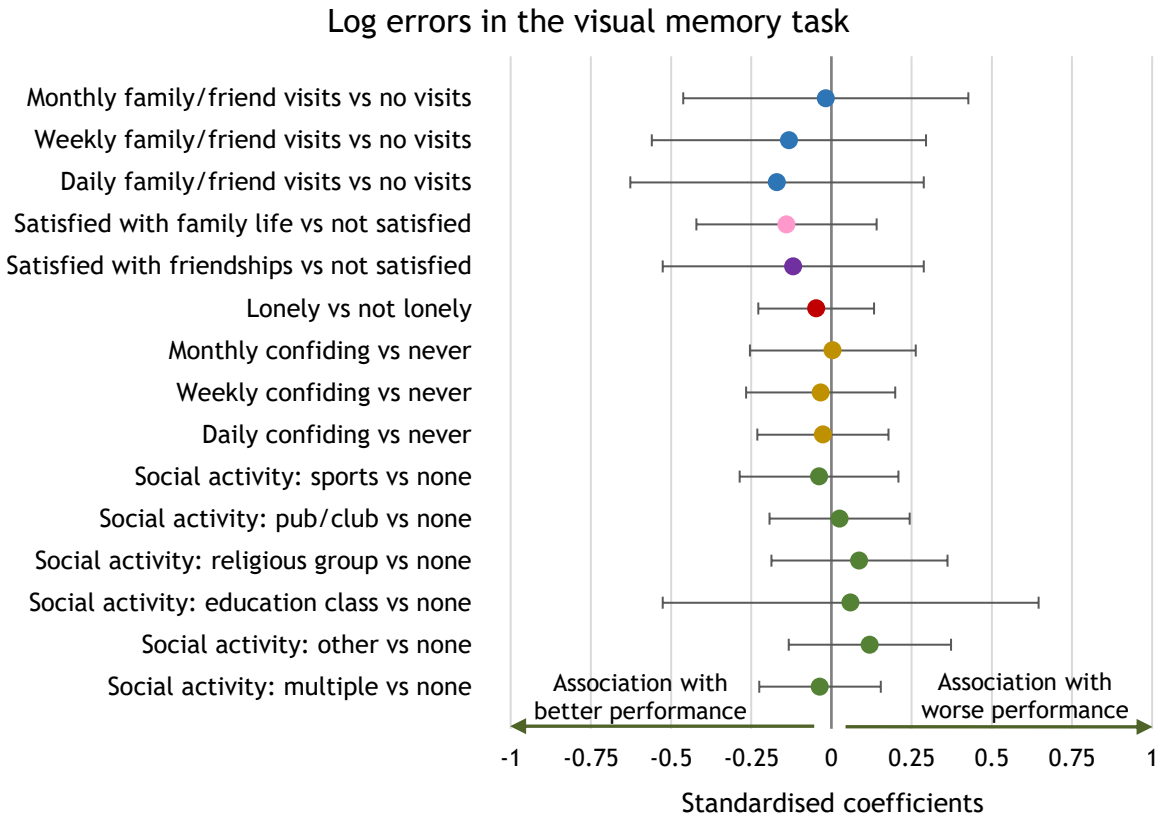


Figure 4-8 Associations of log errors in the visual memory task with proxies of social engagement in a complete model, with 99.7% CI.

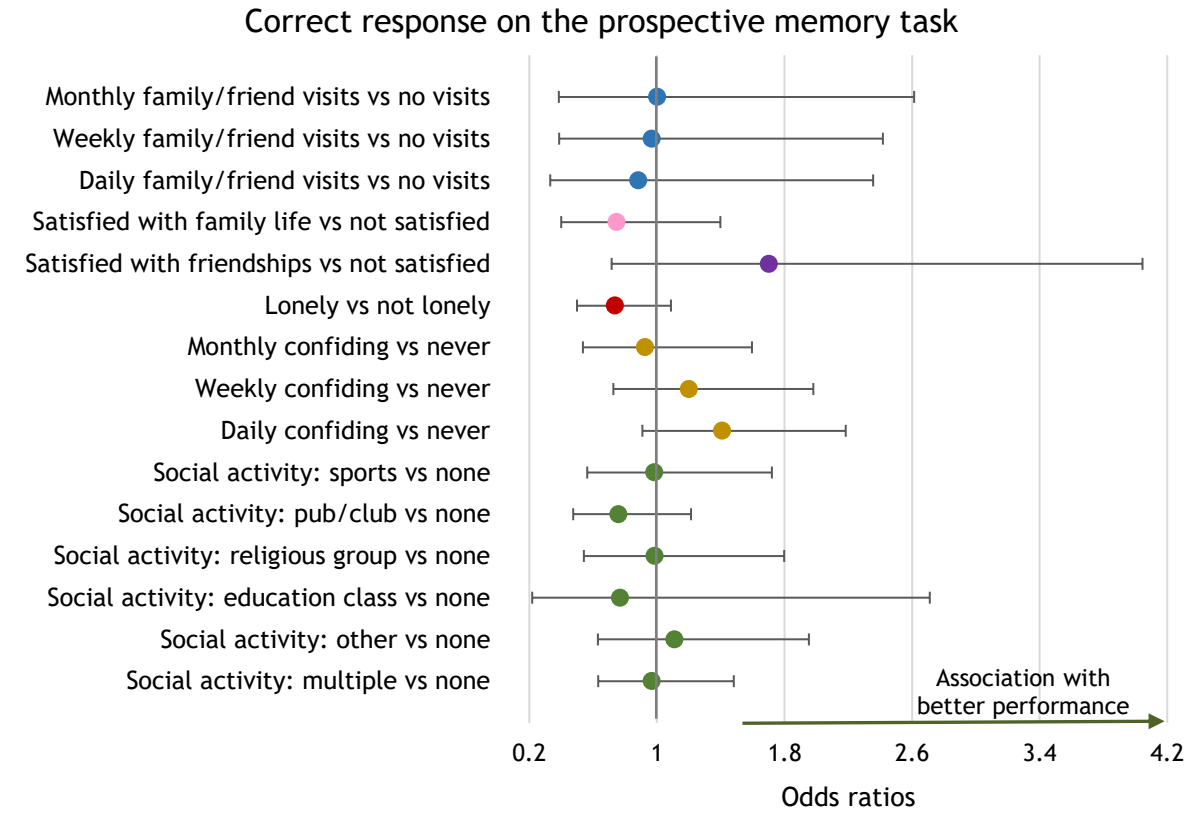


Figure 4-9 Associations of a correct response on the prospective memory task with proxies of social engagement in a complete model, with 99.7% CI.

4.3.4 Study B discussion

Based on cumulative findings from studies involving animal models, non-stroke populations and stroke survivors, I hypothesised that proxies of social engagement would predict post-stroke cognitive performance. My findings partially supported this assumption, indicating relatively modest task-specific associations for some predictors of interest, not all of which remained significant after adjusting for covariates. The experience of loneliness was the only proxy of social engagement to be associated with most tasks, consistently predicting poorer performance.

Many studies have reported an association between loneliness and health-related outcomes, independent even of objective measures of social engagement (297-300). Similar findings have been reported in relation to cognitive function (301). Longitudinal studies in the general older adult population indicated that loneliness is associated with an increased risk of incident dementia and cognitive decline, and an increased rate of the latter (302, 303).

Considering loneliness in a broader sense, as a subjective measure of social engagement, it also seems relevant that perceived social support has been previously identified as a predictor of better performance on word recall and a measure of executive function (304). The cognitive abilities required for completion of these tasks were plausibly similar as for the prospective memory and verbal-numerical reasoning tests here, performance on which was inversely associated with loneliness.

Finding that a subjective measure of social engagement predicted performance on a memory-based task is also the one similarity between my results and those obtained in a previous study investigating cognitive outcomes among stroke survivors, described in the Introduction (245). Beyond that, a comparison of observations is rendered difficult by the difference in chosen social engagement proxies and cognitive tasks. Despite this, there is one conclusion that may be particularly relevant to my findings. The authors reported that social ties and emotional support were positively associated with a cognitive summary score, yet when individual tasks were considered, only performance on one of seven (word recall) was predicted by social factors.

In view of this, it seems possible that I would have found significant associations for more proxies of social engagement or observed greater effect sizes if I had used a global measure of cognitive function. Considering my results on a task-specific level, it is also noteworthy that none of my predictors of interest were associated with visual memory in any of the models. This is consistent with findings in Study A, where the emerging pattern of results appeared to diverge for this particular task, indicating a possible source in psychometric properties.

Regarding reaction time, evidence on its predictors seems to be overall inconclusive, with some studies supporting my findings (305, 306), and others – not (304, 307). Perhaps most surprising is the observed association between religious activity and slower reaction times, with a number of existing reports suggesting a positive relationship between religious engagement and cognition (308). An accurate interpretation of this discrepancy seems however to exceed the scope of my analysis, with potentially multiple factors contributing to the obtained results (309), e.g. relating to used measures of cognitive performance and religious involvement, uncontrolled relevant variables, and residual confounding from included covariates.

In addition to unexpected associations, what seems interesting is that the findings did not support certain expected ones. Specifically, in adjusted models, participation in social activities did not predict better performance on any of the cognitive tasks, despite favourable associations being relatively well-documented at least for sports and adult education (310-314). Regarding the former, the only finding related to faster reaction times in an unadjusted model (consistently with results of Study A), which may reflect that the effect was driven by younger and healthier participants engaging in sports. Thus, once these factors were controlled for, the association was no longer significant. Moreover, the “sports” category would not account for non-occupational physical activities that could be more commonly engaged in by this population, such as gardening or walking. In relation to adult education, on the other hand, it may be relevant that very few stroke survivors reported participation in such an activity (under 2%).

There are also several other explanations to consider, which apply to my findings more broadly. Firstly, many assumptions arise from studies in the general

population, and not all may be directly transferable to the context of stroke. Some proxies of social engagement could also be specifically associated with cognitive change over time, which given the cross-sectional nature of my data, I was unable to investigate. The complexity of associations between social engagement and cognition may further exceed the limits of regression analysis, with some variables potentially mediating the effects of others (298, 315). Moreover, my finding that no proxies of social engagement significantly predicted cognitive performance when combined into one, complete model, may indicate model overfitting.

4.3.4.1 Clinical implications

From a person-centred perspective, a particularly concerning finding was that a quarter of my sample experienced loneliness. Taking into account the risk of participation bias, as well as a recent report of national survey data (316), this is likely an underestimation of true prevalence in an unselected stroke population. Although a negative impact on cognition requires further confirmation, it seems clear that loneliness is a common problem that can severely compromise an individual's well-being. In recognition of this, in 2018 a United Kingdom government press release announced a planned £20 million investment to help socially isolated and lonely people (317).

Evidence from studies conducted in the general adult population suggest that interventions for alleviating feelings of loneliness may indeed be successful (318, 319), yet plausibly there are unique aspects to the experience of loneliness in stroke, as well as particular considerations regarding intervention delivery (e.g. overcoming communication difficulties). One research team has indicated plans to trial an intervention – LISTEN (Loneliness Intervention using Story Theory to Enhance Nursing-sensitive outcomes) – with stroke survivors, however, result of this study seem so far unpublished (320, 321).

Until bespoke interventions for diminishing loneliness following stroke are available, potential avenues of support may be sought within existing resources. A first step would be to identify stroke survivors who experience loneliness – asking even just a single question, which could be done at any point of contact with health services. If flagged as an important issue to the individual, at its

basic level, provision of help could simply involve informing about local befriending services and peer support groups, as contact with people who have similar experiences seems to be of particular merit (322, 323).

4.3.4.2 Research implications

The value of including loneliness as a candidate predictor in prognostic models seems so far to have gone unrecognised; while my findings, combined with the reviewed literature, suggest this factor could indeed contribute to explaining variance in post-stroke outcomes, independently of demographics, health status, lifestyle factors and depression. There are also practical advantages worth considering. Loneliness is a variable that can be easily assessed and despite being subjective, or perhaps because of it, its meaning seems relatively unambiguous.

In comparison, interpreting what objective proxies of social engagement truly represent can be more challenging. Regarding frequency of family/friend visits, unlike those monthly and weekly, I found that daily visits did not predict faster reaction times compared to having no visits in any of the models. In part, this could be due to receiving frequent visits as a result of an individual's greater need for external assistance with activities of daily living. In turn, more frequent visits may reflect less/no support within the household and/or greater functional and cognitive difficulties.

These findings further prompt the notion that more is not always better, as social interactions can also be negative. Opposite to the proposed stress-buffering role of social support, negative social interaction can itself be a source of stress, including e.g. experiences of hostility, discouragement, shaming, deceit or violation of boundaries; and has been linked to an increased risk of disease (324, 325). Moreover, in the specific context of post-stroke recovery, even support can have its negative aspects, as extending to overprotectiveness, it may limit opportunities for stroke survivors to engage in certain activities (326-328). These combined arguments indicate the importance of applying measures that account for the nature of social interactions, in addition to their frequency (e.g. how often do you do something enjoyable with a relative/friend?).

4.3.4.3 Limitations

Many of the limitations discussed for Study A also apply here: employing a cross-sectional design, identification of eligible participants based on self-reported medical history, concerns regarding representativeness of the study sample, high volume of missing data, use of bespoke measures of cognitive function, and inability to control for certain relevant variables (particularly stroke-related factors). A further limitation, unique to this study, relates to a lack of information on basic features of the participants' social networks, specifically, their marital status and number of people they live with. Unfortunately, access to these variables was not covered under the current UK Biobank application.

4.3.4.4 Strengths

My study is one of the few to focus on proxies of social engagement as predictors of post-stroke cognitive performance. Similarly as in the case of Study A, important strengths lie in the relatively large sample size and opportunity to control for multiple variables, plausibly associated with both social engagement and cognitive performance. Regarding the latter, history of depression constituted a key covariate, accounting for which allowed to demonstrate that the effect of some aspects of poor social integration likely extend beyond a deleterious association with mood. A further strength relates to investigating different types of social engagement proxies, representing both ends of the objective-subjective continuum.

4.3.4.5 Future directions

The associations between post-stroke cognition and social engagement are undoubtedly complex, likely involving bidirectional effects. Our understanding of them can be improved through longitudinal studies that not only assess the key predictors and outcome of interest, but also factors that may explain the mechanisms underlying these associations, such as differences in physiological stress markers or likelihood of engaging in behaviours that impact health. However, it is interventional studies that are ultimately needed to demonstrate whether improving aspects of social engagement translates to improved post-stroke cognitive outcome.

As discussed above, the transferability of interventions that were developed in the general population to this context may be poor. Thus, before the stage of bespoke intervention development and implementation is reached, it is important to identify what factors contribute to limited social engagement following stroke and which of these stroke survivors find important to address (as some may result from personal preference). This in turn emphasizes the need for qualitative studies. The few available publications on this topic seem to prove how informative such insights can be, having identified issues around: loss of shared activities with friends, being unable to contribute to one's environment, communication barriers, embarrassment around disabilities, and lack of access to transportation (322, 329, 330).

4.4 Summary

Physical activity, sedentary behaviour and social engagement have so far not been in the focus of prognostic research in stroke. It is unknown whether the accuracy of existing prediction models for post-stroke cognitive outcomes would meaningfully improve after adding such predictors, particularly given their close link to demographics and health-related factors, which are already commonly incorporated. Regardless, the value of continuing research in this direction seems to predominantly lie in identifying novel targets for intervention. Modifying aspects of habitual activity patterns and social engagement could potentially be achieved through means feasible and acceptable to stroke survivors, and improve more outcomes following stroke than just cognition.

Although I conducted two independent studies, this was not in support of an outlook that these factors of interest should be addressed separately. On the contrary, there is evidence indicating that targeting them simultaneously could be of particular benefit. A systematic review of studies in stroke populations concluded that social support is an important motivator for engaging in physical activity (234); while in the general population, findings suggest a bidirectional relationship, where loneliness may reduce the probability of being physically active, and physical activity decreases feelings of loneliness (331).

Regarding which specific aspects of physical activity, sedentary behaviour and social engagement are important to target, my most consistent observations

indicated that cognitive performance was positively associated with computer use (a mentally active sedentary behaviour), and negatively associated with watching TV (a mentally passive behaviour) and experiencing loneliness. However, given multiple study limitations, these results should be viewed as hypothesis-generating rather than definitive. Unfortunately, until bespoke studies addressing these topics are conducted, future research is likely to face similar challenges.

Large population-based general-purpose cohort studies have many key strengths. However, when focusing on a particular condition (such as stroke, here), the available sample size considerably decreases and information specific to that condition is often lacking, and thus cannot be accounted for. Conversely, datasets from clinical sources are likely to include such information and also offer greater certainty regarding sample representativeness. Yet, apart from data on commonly recognised risk factors, such as smoking and alcohol intake, little or no insight is provided regarding individuals' everyday life activities and experiences.

Fortunately, this does not imply that existing clinical datasets have no application to investigating potentially modifiable predictors of post-stroke cognition. Given that: i) at least part of the effect of lifestyle factors and life experiences on cognitive impairment and decline is driven by detrimental changes to health, and ii) some diseases can be alleviated or even reversed; one avenue for using clinical data to its best advantage may be through determining the associations between post-stroke cognition and comorbid conditions.

Chapter 5 Cardiovascular risk factors as predictors of acute post-stroke cognitive function: Are there two sides to this story?

In the previous chapter, I presented two studies investigating the relevance of potentially modifiable factors to cognitive function following stroke, with a specific focus on habitual physical activity, sedentary behaviour, and social engagement. In the section on *Routes to affecting cognition* that formed part of my introduction, I indicated that the impact of these factors is considered as at least partially driven by modifying cardiovascular risk burden (particularly in relation to physical activity patterns). Interestingly, however, findings from observational studies and trials of interventions to alleviate cardiovascular diseases are not consistent in supporting a link between variation in such risk factors and post-stroke cognitive outcomes.

In this chapter, I argue that (to some extent) neutral results could be due to the complexity of these associations, which may involve differing paths of influence and interactions between comorbid conditions. To test my assumptions, I developed a moderated mediation model, using data from a hyper-acute stroke unit setting. This chapter is an adaptation of my published work (332), with edited and expanded Methods, Results and Discussion sections.

5.1 Introduction

Many studies have investigated the effects of prevalent cardiovascular risk factors on post-stroke cognition (23, 25, 333). This interest seems unsurprising given the high comorbidity burden among stroke survivors, encompassing conditions recognised as predictors of age-related cognitive decline and dementia, such as: diabetes mellitus, hypertension, coronary and peripheral vascular disease, atrial fibrillation, and previous stroke (334-337). Interestingly, despite a strong premise to assume the relevance of cardiovascular risk factors to post-stroke cognition, only two of eleven prognostic rules described in Chapter 3 included a predictor of this type. In part, this could be due to actual

evidence on the associations between comorbid conditions and post-stroke cognitive function appearing inconclusive or conflicting (e.g. see 23, 145).

One possible explanation for the inconsistency of observations relates to the complexity of this relationship. Cardiovascular conditions often pre-date incident stroke and so it seems plausible that to some degree post-stroke cognitive impairment is a manifestation of precursory vascular neurodegenerative processes (338, 339). However, the effect of cardiovascular risk factors on post-stroke cognition is likely not only driven by gradual neurodegeneration. Certain conditions are associated with stroke severity, which is in turn a major determinant of cognitive outcome.

Intuitively, it seems that the presence of cardiovascular diseases should be consistently detrimental across outcomes. Indeed, this seems to be the case for atrial fibrillation, which is associated with both higher incidence of dementia and the most severe ischaemic stroke subtype – cardioembolic infarction (17, 340, 341). However, the effects of other cardiovascular risk factors on cognitive function may be more equivocal, particularly where the pathophysiological processes they contribute to trigger endogenous adaptive mechanisms. For example, transient ischaemia has been reported to induce a state of “ischaemic tolerance” or “preconditioning“ that temporarily protects tissue from subsequent, persistent ischaemia (342). Evidence from observational studies suggests that this phenomenon, which is consistently demonstrated in animal models, may also occur in clinical practice. In cases of stroke, prodromal TIA has been associated with less severe symptoms, smaller infarct volumes, and better functional outcomes (343-345).

A similar example relates to vascular disease. Its most common forms involve build-up of atherosclerotic plaque, leading to narrowing of vessels and thus reduction of blood flow (346). Although this is a progressive pathological process, it may support the advantageous development of collateral circulation. In the event of arterial occlusion, robust collaterals sustain perfusion, helping to maintain nutritive tissue needs (347, 348). Following stroke specifically, efficient collateral circulation has been associated with favourable clinical outcomes, including an improved response to thrombolytic and recanalisation therapy (349, 350).

To add further complexity, manifestation of the described putative protective mechanisms may be affected by comorbidity. Observations from clinical and preclinical models suggest that both hypertension and diabetes may impair the development of collaterals (346, 351-355), while diabetes also precludes ischaemic tolerance (356, 357). These findings, therefore, highlight that cardiovascular risk factors may not only co-occur but also interact.

In this context, it seems that the traditional approach to data analysis, involving use of multivariable regression models, may have been insufficient to capture the true nature of associations between cardiovascular diseases and post-stroke cognition. In its basic form, the method can only identify those factors directly associated with an outcome, while remaining factors are held constant. As such, it does not allow us to explore the potential for multiple routes of predictor impact, nor the interaction between co-occurring diseases. It is therefore possible that neutral results, reported from some of the previous studies, stem from the duality or conditionality of considered effects.

In my study, I aimed to investigate how cardiovascular risk factors can affect cognitive function in the acute phase after stroke, through influence on stroke severity and prior cognitive impairment. I specifically hypothesised that:

- previous TIA and vascular disease may have differential effects on acute post-stroke cognitive function depending on the path of influence, possibly predicting poorer performance through an increased risk of prevalent dementia, while predicting better performance through an association with reduced stroke severity;

- a favourable effect of vascular disease on stroke severity, and in turn on acute cognitive performance, may be conditional on the absence of hypertension and diabetes;

- a favourable effect of previous TIA on stroke severity, and in turn on acute cognitive performance, may be conditional on the absence of diabetes.

5.2 Methods

The dataset used for my analyses is part of a larger research database. The West of Scotland Research Ethics Committee approved the primary project on the 4th of February 2016 (16/WS/0001). As data collection was embedded in routine clinical care, and patient information was fully anonymised prior to archiving, obtaining written informed consent from participants was not required. I based the design and conduct of the present study on recommendations from recent works, summarising theoretical and practical approaches to development of mediation and moderation models with an emphasis on best practice (358, 359).

5.2.1 Study setting and participants

Participants were consecutive patients admitted to the hyper-acute stroke unit of Glasgow Royal Infirmary. The unit provides high dependency level clinical care, accepting all cases of suspected stroke and TIA, regardless of preadmission physical and cognitive function. Collection of anonymised data took place in four waves: May 2016 to February 2017, April to June 2017, October to December 2017, and July to August 2018. For the purposes of this study, I excluded patients for whom a diagnosis of stroke or TIA had been ruled out by the clinical team.

5.2.2 Data collection

Together with four other trained researchers, we used medical records and data collected by the clinical team during acute admission to extract information on basic demographics, pre-existing medical conditions, and findings from neurological examinations. We additionally acquired cognitive data through directly administering a cognitive screening test, which I describe in detail in a subsequent section.

5.2.2.1 Predictors

Based on indications from previous research, I included cardiovascular risk factors with a plausible association with post-stroke cognitive function: vascular disease (peripheral and coronary), atrial fibrillation, hypertension, diabetes mellitus, previous stroke, and previous TIA. The clinical process in the stroke service is for these data to be confirmed from at least two sources, and includes

information from both primary and secondary care records. I coded all risk factors as either present or absent. I additionally accounted for basic demographics of sex and age, the latter treated as a continuous variable (years).

5.2.2.2 Mediators

The first mediator I included in the model was stroke severity, assessed using the NIHSS (71, 72). Where a specific NIHSS score had not been documented by the clinical team, the researcher responsible for data collection would retrospectively derive a score based on findings from acute neurological examinations, described in patient notes (360). As per emergency department triage policy, examination is performed immediately upon hospital admission and then confirmed in the hyper-acute stroke unit, noting any changes in initial signs (resolution or progression). For inclusion in the analysis, I categorised NIHSS into four groups: no stroke signs (score of 0), minor stroke (score of 1 to 4), moderate stroke (score of 5 to 15), and severe stroke (score of 16 to 42) (361).

The second mediator was a formal diagnosis of dementia prior to incident stroke or TIA. This information was obtained from primary or secondary care medical records, including reports from mental health services. In the United Kingdom, dementia is diagnosed by specialist (secondary) care providers, based on the International Classification of Diseases (ICD) criteria (362).

5.2.2.3 Cognitive performance

Our research team assessed cognitive performance within a week of stroke or TIA, using a test battery of 13 items, comprising Hodkinson's Abbreviated Mental Test (AMT-10) (363, 364) and a short-form version of the MoCA (77). The following tasks were included: stating one's age, current time to the nearest hour, date, place, recognition of two people, date of birth, year World War I began, name of current Prime Minister, months of the year listed in a backwards order, five-word recall, clock draw, recent news item, and verbal fluency (words beginning with "f"). Clock draw was the only task to be scored out of 3 points (face, number, hands), with remaining items assigned a single point, summing up to a maximum total of 15.

For the purpose of this study, I considered outcome data to be missing under three conditions: (i) the patient refused to participate; (ii) the patient was discharged prior to assessment; (iii) the assessment was initiated but could not be completed due to external circumstances, e.g. to avoid disruption to work carried out by the clinical team. If a patient was unable to complete a particular task due to an existing impairment (e.g. aphasia or limb weakness), I assigned a score of zero for that item, including it in the sum score (365).

In cases where the severity of the patient's condition (e.g. altered level of consciousness, agitation) precluded from attempting the assessment altogether, I entered a total score of -1. This approach, where an untestable status is assigned the lowest possible score, mimics the scoring system applied in delirium screening (66, 67). It allowed me to minimise missing data and avoid exclusion of participants with the most severe presentation. For inclusion in the analysis, I divided cognitive scores of all participants into quintiles, creating the following groups: (i) scores from -1 to 2, (ii) scores from 3 to 8, (iii) scores from 9 to 11, (iv) scores of 12 and 13, and (v) scores of 14 and 15.

5.2.3 Statistical analysis

I developed a first stage dual moderated mediation model for prediction of cognitive performance, with two parallel mediators – stroke severity and previous diagnosis of dementia (Figure 5-1). Although it was possible to model a number of different moderation effects, I focused on three interactions most consistently demonstrated by existing evidence (346, 351-357). Namely, I hypothesised that mediation of the effect of vascular disease and previous TIA on the outcome through stroke severity may be moderated by the presence of diabetes (in both cases) and hypertension (for vascular disease only).

Given that mediation analyses assume causal relationships, I aimed to build a model the structure of which would reflect the actual temporal order of occurrences. This order was definite for paths mediated by stroke severity (with cardiovascular risk factors present before the index stroke/TIA, and the cognitive assessment taking place after) and appeared plausible for paths mediated by dementia, with evidence suggesting that cardiovascular diseases would have likely developed in earlier stages of life.

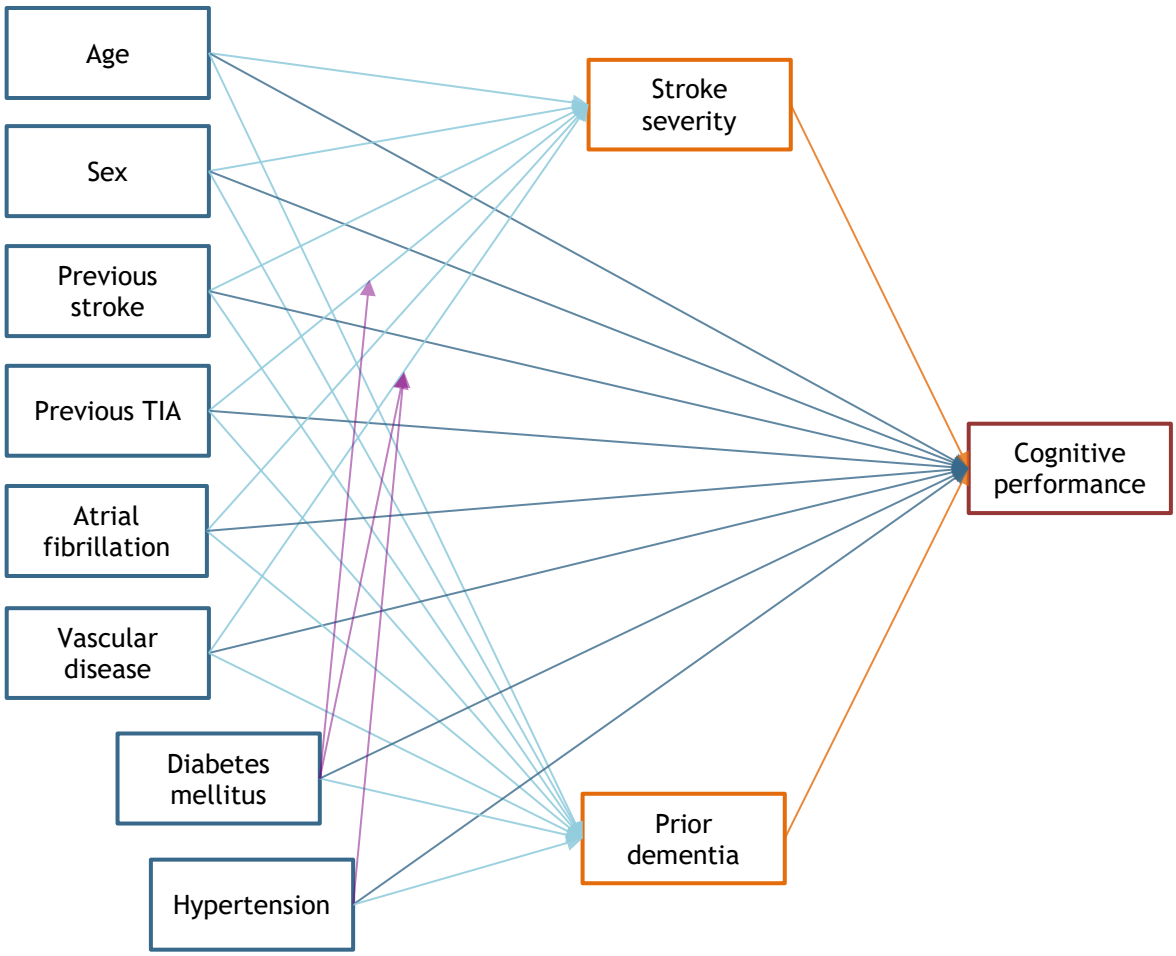


Figure 5-1 Conceptual diagram of the proposed dual moderated mediation model with two parallel mediators for acute cognitive performance.

I analysed the data within a path analysis framework, using structural equation modelling (SEM) software, Mplus version 8.3 (366). As the outcome of interest was an ordinal variable, I estimated associations with predictors and mediators based on a probit regression, using a robust weighted least squares mean and variance-adjusted estimator (WLSMV). Missing data were handled as per software default, that is, cases with missing data on predictors were removed from the analysis, while missing outcome and mediator data were estimated as a function of the observed predictors (367).

The moderated mediation analysis involved regressing cognitive performance on both mediators and the eight predictors, while regressing each mediator on the eight predictors. In line with my hypothesis, I also regressed stroke severity on three interaction terms (TIA x diabetes mellitus, vascular disease x diabetes mellitus, vascular disease x hypertension). As I based decisions regarding inclusion of variables in the analysis on research evidence, in order to avoid

model overfitting, I intended to retain all predictors and both mediators, regardless of path significance. However, in order to achieve a more parsimonious model, I planned to remove nonsignificant interaction terms (368). I used estimates obtained through the regression analysis to calculate indirect effects, applying a product of coefficients approach (359, 369).

For significant interaction terms, I quantified the indices of partial mediated moderation. Using the more complex case of vascular disease as an example, this entailed estimating how much the mediated effect of this factor on cognitive performance changed depending on the following: firstly, the presence or absence of diabetes mellitus, assuming absence of hypertension (held fixed, as all other predictors); secondly, the presence or absence of hypertension, assuming absence of diabetes mellitus (370). I then probed the partial moderated mediation effects to establish for what specific combination of factors (four options based on presence vs absence of diabetes and hypertension) vascular disease had a significant conditional indirect effect on cognitive performance. Based on the same principles, I planned to apply a simplified version of this procedure to estimate the conditional indirect effect of TIA, depending on the presence or absence of diabetes.

I determined the significance of individual paths and indirect effects through constructing bias-corrected bootstrap confidence intervals, based on drawing 1000 bootstrap samples. This method is recommended as one that does not assume normal sampling distribution and offers greater precision for calculating confidence intervals compared to alternatives (371, 372). There are currently no consensus guidelines regarding sample size requirements in SEM. However, previous simulation studies applying bootstrapping have determined that to detect small mediation and moderated mediation effects (estimate = 0.14) with 80% power, a sample of nearly 500 participants is required (373, 374).

To provide information on the magnitude of mediated effects, I calculated the proportion-mediated effect size – a ratio of the specific indirect effect to the total effect of a predictor (375). This is considered an intuitive measure and is easily extrapolated from a simple to a multi-mediator model (358). Expecting that the direct and indirect effects of a single predictor may be of opposite signs, I planned to use absolute coefficient values (376).

5.2.4 Additional analyses

Recognising the potential bias from assumptions about missing data, I repeated the described procedures in a sensitivity analysis, using a more conservative approach. Namely, I excluded participants who due to an existing impairment did not complete particular tasks within the cognitive assessment.

Further, to examine how use of different analysis strategies may impact findings, I additionally conducted a basic multivariable regression, as commonly employed in previous studies. Specifically, I entered stroke severity and prior dementia into the model alongside all other predictors, thus regressing cognitive performance on a total of 10 variables (Figure 5-2).

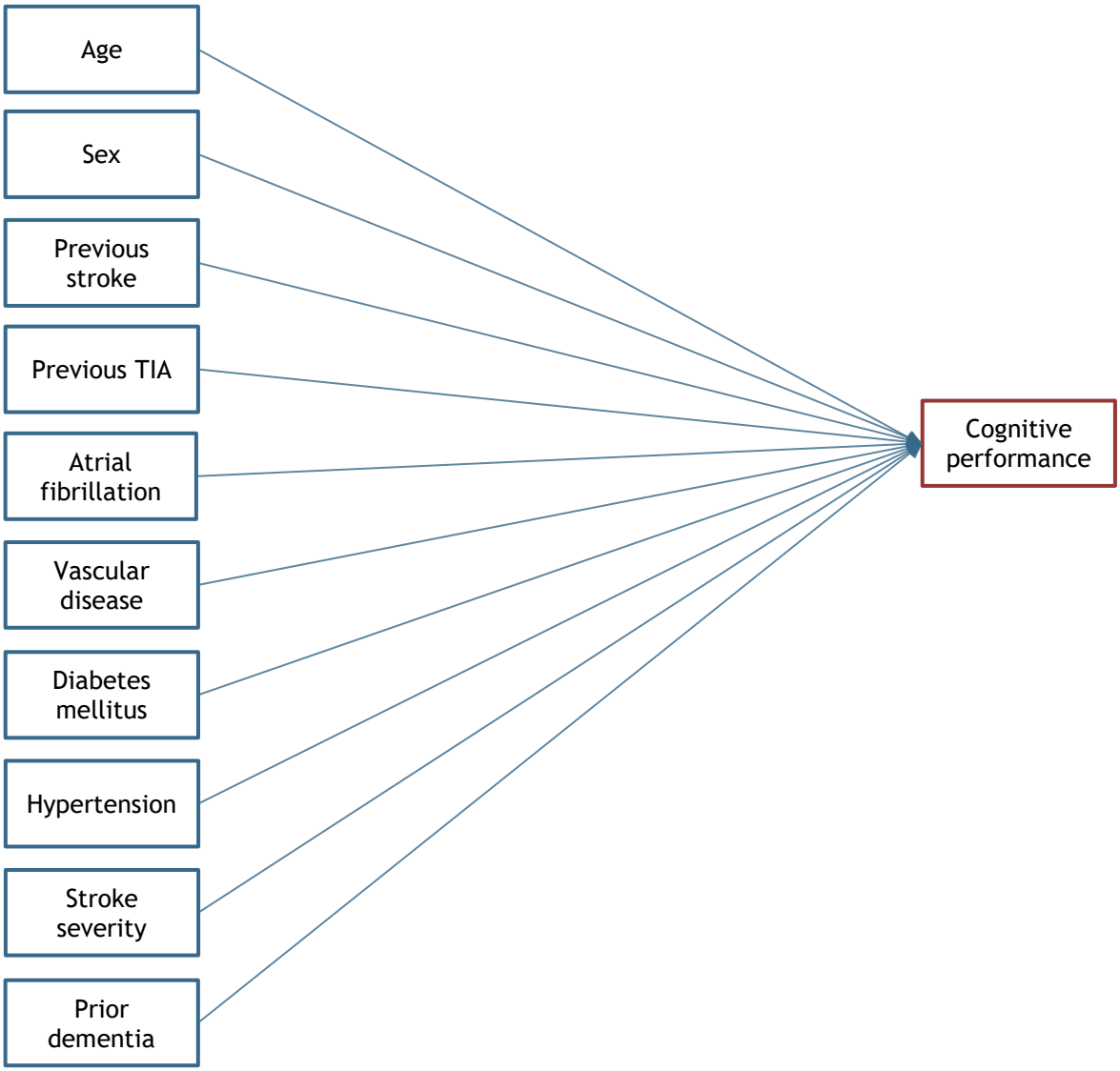


Figure 5-2 Conceptual diagram of a basic multivariable regression model with acute cognitive performance regressed on 10 predictors.

5.3 Results

A total of 703 patients were screened during the study recruitment waves. From this sample, 109 were given a final diagnosis other than stroke or TIA, leaving 594 participants fulfilling my inclusion criteria. Table 5-1 presents characteristics of the study sample. A correlation matrix for variables incorporated in the model is provided in Appendix 6. As seven patients had missing data on predictor variables, 587 participants were included in the final analysis.

Table 5-1 Characteristics of study sample.

Variables	
Age (years)	
Range	26 to 100
Median (IQR)	72.0 (21.0)
Missing	2
Sex (female)	
N (%)	269 (45.3%)
Missing	0
Previous stroke	
N (%)	136 (22.9%)
Missing	0
Previous TIA	
N (%)	40 (6.7%)
Missing	0
Atrial fibrillation	
N (%)	108 (18.2%)
Missing	5
Diabetes mellitus	
N (%)	124 (20.9%)
Missing	5

Table 5-1 Baseline characteristics of study sample. *Continued*

Variables	
Hypertension	
N (%)	316 (53.2%)
Missing	5
Vascular disease	
N (%)	149 (25.1%)
Missing	5
Prior dementia	
N (%)	57 (9.6%)
Missing	0
Stroke severity (NIHSS score, range: 0 – 42)	
Range for sample	0 to 31
Median (IQR)	3.0 (1 – 5)
Categories	
No stroke signs, N (%)	93 (16.1%)
Mild, N (%)	321 (55.4%)
Moderate, N (%)	128 (22.1%)
Severe, N (%)	37 (6.4%)
Missing	15
Cognitive test score (range: 0 – 15)	
Range for testable participants	0 to 15
Median for testable participants (IQR)	11.0 (8 – 13)
Untestable participants, N (%)	101 (17.0%)
Missing	22

IQR indicates interquartile range; NIHSS, National Institutes of Health Stroke Scale; TIA, transient ischaemic attack.

5.3.1 Final model structure and properties

As described in the Methods section, the initial model included three interaction terms. However, preliminary results indicated that the interaction term of TIA and diabetes mellitus was not significantly associated with stroke severity ($p = 0.560$) and, therefore, I removed it from the model. Subsequent findings suggested a trend for the remaining two interaction terms, between vascular disease and diabetes mellitus ($p = 0.057$) and vascular disease and hypertension ($p = 0.056$), and so I opted to retain them. Consequently, the final model differed from that presented in Figure 5-1 in only one aspect, namely, I did not consider diabetes as a moderator for the effect of TIA on stroke severity (Figure 5-3).

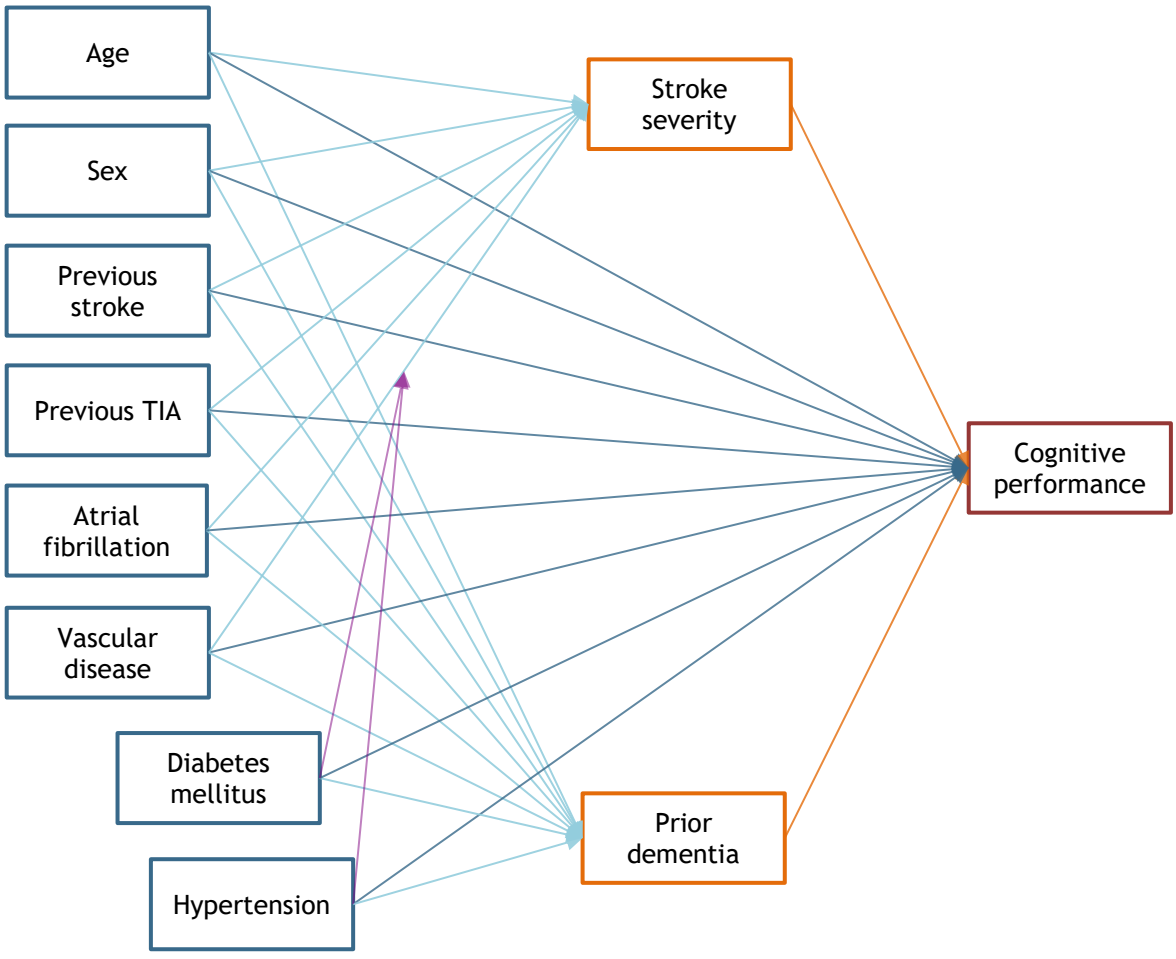


Figure 5-3 Conceptual diagram of the final moderated mediation model with two parallel mediators for acute cognitive performance.

For this model, the chi-square statistic indicated no significant discrepancy between the observed and model-estimated covariance matrices: $X^2 = 6.580$, $p = 0.254$. Additional recommended fit indices confirmed good model fit: Root

Mean Square Error of Approximation (RMSEA) = 0.023, Comparative Fit Index (CFI) = 0.995, and Standardised Root Mean Square Residual (SRMR) = 0.030 (377). Overall, my model explained R^2 = 62.10% of variance in cognitive test scores.

5.3.2 Associations between predictors and mediators

More severe strokes were associated with age and atrial fibrillation, while severity decreased with a history of previous TIA (Table 5-2). The observed associations with interactions terms, just above the threshold of statistical significance, suggested opposing effects of vascular disease, depending on the presence of diabetes mellitus and hypertension. Co-occurring with the former, it appeared potentially associated with greater stroke severity, while co-occurring with the latter – with less severe presentation. Predictors of prior dementia included age and previous stroke. I also observed a trend for an association between dementia and history of vascular disease (p = 0.054).

Table 5-2 Direct associations between predictors and stroke severity, dementia and cognitive performance.

	Unstandardised coefficients (95% bias-corrected CI)		
	Stroke severity	**Prior dementia**	**Cognitive performance**
Age	0.012 (0.004, 0.019)*	0.059 (0.039, 0.078)*	0.003 (-0.016, 0.027)
Sex (female)	-0.031 (-0.210, 0.180)	-0.098 (-0.458, 0.248)	-0.264 (-0.588, 0.046)
Previous stroke	0.008 (-0.217, 0.222)	0.538 (0.176, 0.932)*	0.230 (-0.129, 0.649)
Previous TIA	-0.512 (-0.934, 0.147)*	-0.342 (-3.886, 0.279)	-0.141 (-2.136, 0.510)
Atrial fibrillation	0.355 (0.075, 0.609)*	0.145 (-0.270, 0.554)	-0.092 (-0.464, 0.279)
Diabetes	-0.025 (-0.274, 0.209)	-0.028 (-0.636, 0.535)	-0.041 (-0.604, 0.508)
Hypertension	0.076 (-0.146, 0.301)	-0.133 (-0.571, 0.377)	-0.065 (-0.502, 0.364)
Vascular disease	0.002 (-0.405, 0.374)	0.611 (-0.051, 1.246)	0.390 (-0.178, 1.127)
Vascular disease x diabetes	0.466 (-0.031, 0.924)	———	———
Vascular disease x hypertension	-0.486 (-0.971, 0.016)	———	———

*Significant at p < 0.05

CI indicates confidence interval; TIA, transient ischaemic attack.

5.3.3 Direct effects on cognitive performance

The results indicated that both mediators were associated with acute post-stroke cognitive performance: coefficient = -0.748, 95% bias-corrected CI: -0.963 to -0.572 for stroke severity; coefficient = -0.720, 95% bias-corrected CI: -1.096 to -0.444 for dementia. However, I observed no significant direct effects of included predictors on cognitive performance.

5.3.4 Indirect effects on cognitive performance

5.3.4.1 Effects mediated through stroke severity

I found that age had a negative specific indirect effect on cognitive performance (Table 5-3), with 16.36% of the absolute overall effect of age on cognition mediated by stroke severity. Poorer cognitive outcome was also indirectly associated with a history of atrial fibrillation, with a proportion-mediated effect size of 57.45%. Conversely, I observed improved cognitive performance through a specific indirect effect of previous TIA, which constituted 49.68% of the absolute overall effect.

The indices of partial mediated moderation suggested a trend for both estimated conditional indirect effects of vascular disease on cognition through stroke severity (p = 0.077 for both diabetes mellitus and hypertension). Through probing, I found that vascular disease produced a significant positive effect on performance under only one condition – where there was a history of hypertension without diabetes mellitus (estimate = 0.362, 95% bias-corrected CI: 0.032 to 0.675; p = 0.024). The proportion-mediated effect size was 30.37%.

5.3.4.2 Effects mediated through prior dementia

In relation to the second mediator, dementia, I observed that previous stroke had a negative specific indirect effect on cognition, constituting 62.12% of the absolute overall effect. Despite the noted trend for an association between dementia and vascular disease, the specific indirect effect of this risk factor on cognitive performance did not reach statistical significance (p = 0.089). Age, therefore, was the only predictor to exert a significant negative indirect effect on cognition through both mediators. Compared to stroke severity, dementia conveyed a considerably larger portion of its overall absolute effect – 78.18%.

Table 5-3 Indirect associations between predictors and cognitive performance.

	Unstandardised coefficients (95% bias-corrected CI)	
	Effects mediated through stroke severity	**Effects mediated through prior dementia**
Age	-0.009 (-0.015, -0.003)*	-0.043 (-0.069, -0.024)*
Sex (female)	0.023 (-0.145, 0.172)	0.071 (-0.186, 0.349)
Previous stroke	-0.006 (-0.175, 0.168)	-0.387 (-0.753, -0.130)*
Previous TIA	0.383 (0.083, 0.745)*	0.247 (-0.218, 2.152)
Atrial fibrillation	-0.266 (-0.493, -0.052)*	-0.105 (-0.479, 0.202)
Diabetes	0.019 (-0.161, 0.207)	0.020 (-0.398, 0.516)
Hypertension	-0.057 (-0.244, 0.109)	0.096 (-0.278, 0.476)
Vascular disease	-0.002 (-0.297, 0.322)	-0.440 (-0.981, 0.064)
Vascular disease x diabetes	-0.349 (-0.748, 0.030)	———
Vascular disease x hypertension	0.363 (-0.032 – 0.759)	———

*Significant at $p < 0.05$

CI indicates confidence interval; TIA, transient ischaemic attack.

5.3.5 Results of additional analyses

For the sensitivity analysis, I excluded 38 participants who due to existing impairments were not able to complete particular cognitive tasks. Estimates of direct and indirect effects are presented in Appendix 7, Supplemental Tables 13 and 14. Overall, the findings were similar to those obtained through the main analysis, with differences specifically relating to associations with dementia. Namely, I found a reversed pattern for dementia predictors, where here the association with vascular disease was statistically significant and with previous stroke – at trend level ($p = 0.056$). Moreover, the indirect effect of previous stroke on cognition did not reach statistical significance ($p = 0.080$).

Results of the alternative, basic regression model resembled the direct associations I observed for the moderated mediation analysis, although here three variables were significantly associated with poorer cognitive performance. These included: stroke severity (estimate = -0.753, 95% bias-corrected CI: -0.875 to -0.626; $p < 0.001$), dementia (estimate = -0.995, 95% bias-corrected CI: -1.321

to -0.685; p < 0.001) and age (estimate = -0.028, 95% bias-corrected CI: -0.037 to -0.021; p < 0.001). The model explained R^2 = 45.10% of variance in cognitive performance.

5.4 Discussion

The results of my study provide evidence for the role of stroke severity and prior cognitive impairment in mediating the effects of specific cardiovascular risk factors on acute cognition. Some of my findings were in line with previously reported associations and seem intuitively correct. Poorer cognitive performance was associated with: atrial fibrillation through increased stroke severity; previous stroke through increased risk of prevalent dementia; and with age through both mediators. Importantly, however, I also obtained results which contradict the concept that comorbidities have consistently adverse effects on outcome. Specifically, I found that previous TIA and vascular disease – considered risk factors for cognitive impairment – may be related to better acute cognitive performance through alleviating stroke severity.

At the same time, not all findings supported my hypotheses. Perhaps most interestingly, I observed that the likely positive effect of vascular diseases on cognition was conditional on the simultaneous absence of diabetes and history of hypertension. I assumed that the latter would be detrimental, with previous studies showing that high acute blood pressure, a state often seen in patients with chronic hypertension (378), is associated with poorer prognosis after stroke (379).

It seems, therefore, that the relationship between hypertension and post-stroke outcomes may be indeed more complex than previously suggested, as demonstrated in a recent clinical study (380). The authors found that in cases of major stroke reperfusion, acute high blood pressure was associated with better collateral flow and thus decreased infarct growth and better clinical outcomes, while the opposite was observed for patients without reperfusion. Yet, in relation to my results, it is important to note that the indices of partial mediated moderation did not reach statistical significance, and so I cannot conclude that there is indeed a difference in the effect of vascular disease on cognition between patients with and without diabetes and hypertension.

5.4.1 Clinical implications

As concluded in Chapter 3, at present none of the existing prognostic rules for post-stroke cognitive impairment can be recommended for clinical use. Nonetheless, the conducted model development studies provide important insight into what factors are most relevant to post-stroke cognitive function. My findings add to this knowledge, highlighting the need to account for comorbidity and the potential for risk factors not only to co-occur, but also interact. Although further confirmation is necessary, it seems plausible that for patients with a history of TIA and vascular disease with hypertension, the risk of cognitive impairment could be underestimated. Such individuals are more likely to present with less severe strokes, while still being prone to the progressive neurodegenerative effects of these conditions, demonstrated in previous studies.

5.4.2 Research implications

The findings indicate that the prevalent use of basic multivariable regression models to determine predictors of post-stroke cognitive function may be overly reductionist. Results of the additional analysis showed that if I were to rely on the same approach, I would not have observed any significant associations between cardiovascular risk factors and acute cognitive performance. Taking into account that effects may differ in direction depending on the path of influence is an important consideration for future studies. The aims would be to improve outcome prognosis and investigate the detrimental role of comorbidity, or the benefits of endogenous adaptive mechanisms and disease management.

5.4.3 Strengths and limitations

The study sample was representative of a real-world stroke population. My method of scoring and coding performance on the cognitive screen allowed to avoid exclusion of patients with the most severe impairments, thus reducing risk of bias. At the same time, I conducted a sensitivity analysis in a subgroup of participants with complete cognitive data to reflect a more conservative approach. Further, in the conduct of this research, I strove to adhere to current best practice guidelines for mediation and moderated mediation analysis.

The data collection protocol, however, was not specifically designed for this work and, in turn, not all relevant information was accessible. This was the case for education – an important covariate to consider, given associations with both cognitive performance and cardiovascular risk factor prevalence and outcomes (381). There was also limited data describing participant index stroke. While I as was able to include stroke severity as an essential component of my model, I could not account for other plausibly relevant features, such as infarct location, volume, or the stroke mechanism (e.g. cardioembolic or small vessel occlusion).

Further, although the idea for the study was inspired by concepts around endogenous adaptations, processes underlying analysed associations cannot be investigated without accounting for additional variables, for example, the extent of cerebral collateral development, or time elapsed between previous TIA and subsequent stroke. However, even with information regarding the latter, it would be very difficult to assess in an observational study whether ischaemic preconditioning indeed occurs (382). While a number of publications have reported that TIA reduces the impact of subsequent ischaemia (as I described in the Introduction to this Chapter), some have found no such association, or even observed a trend toward greater disability following strokes preceded by TIA (382-384). Individual heterogeneity among stroke survivors, and the specific aetiology of cerebral infarction, are likely to influence the apparent relationship between prior TIA and short and long-term post-stroke outcomes (385).

In view of the above, it is important to also consider alternative explanations for the effects of TIA and vascular disease on alleviating stroke severity, which I observed in my study. The role of treatment, which I was not able to control for, seems of particular relevance here. Specifically, research findings suggest that aspirin, routinely administered following TIA, reduces the severity of early subsequent stroke (386), while statins, prescribed in cases of vascular disease, enhance collateral circulation (387, 388).

Further study limitations relate to a retrospective assessment of risk factors from medical case records, as it is possible that relevant conditions had not been noted or even diagnosed. Of particular concern is correct identification of prior TIA cases. Low public awareness of TIA symptoms and significance (253), coupled with a transient nature, may result in omitting to seek professional help, and so

having no record of the event. Multiple factors may have also contributed to underdiagnosis of dementia, which according to research evidence is a common issue (389); while milder forms of cognitive impairment prior to incident stroke were not captured through data collection at all. Dichotomisation based on a formal diagnosis of dementia imposes an assumption that people are either cognitively intact or have a severe form of cognitive impairment, which does not reflect the true, gradual nature of cognitive deterioration.

Finally, the focus and thus conclusions of this study are restricted to acute cognitive outcome. In this context, it is important to note that longitudinal studies have demonstrated considerable individual changes in cognitive status between the acute and chronic stages following stroke (29, 390, 391). Nonetheless, early post-stroke cognitive impairment has been shown to be a predictor of future outcomes, both cognitive and functional (392). Moreover, in healthcare settings, for many stroke survivors, the only opportunity to undergo a cognitive screen may be during hospital admission.

5.4.4 Future directions

Ideas and the motivation for future research on this topic largely stem from the study limitations discussed above. A better understanding of the associations between cardiovascular diseases and post-stroke cognition could be achieved by accounting for variables relevant to the presence of pathology-driven protective adaptations and treatment effects. Capturing milder forms of prior cognitive impairment, on the other hand, could render more precise estimates for the strength of mediated effects, which here could have been underestimated.

Further, it seems important to explore how the role of comorbidity in shaping cognitive outcomes may differ across time. It is possible that some factors, for which I found no evidence of an unconditional, independent effect on acute cognition (e.g. diabetes), are more relevant to longer-term outcome. Similarly, observed associations might vary depending on whether cognitive function is considered at a single timepoint or in terms of changes in status over time. It is these issues that will be the focus of my final study.

5.5 Summary

In this study, I found that the effects of specific cardiovascular risk factors on acute post-stroke cognitive function are partially mediated through stroke severity and prior dementia. Not all of these effects were detrimental. Vascular disease, conditional on the presence of hypertension and absence of diabetes, and previous TIA seemed associated with better cognitive performance through reduced stroke severity. My observations highlight the complexity of associations between cardiovascular risk factors and post-stroke cognition. In this context, basic multivariable regression models may be overly reductionist, leading to the misidentification of important, potentially casual relationships.

Chapter 6 The Assessing Post-stroke Psychology Longitudinal Evaluation (APPLE) study: Design, participants, and data collection

In light of limitations affecting my two previous studies, I have mentioned the important advantages the availability of a bespoke stroke cohort, with longitudinal cognitive follow-up, could have to progressing prognosis research on post-stroke cognitive change. The next three chapters of my thesis are founded in my opportunity to contribute to developing such a resource – the Assessing Post-stroke Psychology Longitudinal Evaluation (APPLE) dataset. From recruiting participants and conducting assessments, to quality control of documents received from multiple, external research sites, to resolving queries generated by an independent clinical statistics service – work on APPLE constituted the single greatest task during my PhD studentship. Before reporting on my use of this dataset to address my specific thesis aims, I give an overview of the APPLE project, with a particular focus on inclusion criteria, the consent process, and participant assessments.

6.1 Key features

APPLE was a multicentre, prospective cohort study, developed with an overarching aim to improve our assessment and understanding of the short and longer-term neuropsychological consequences of stroke. Specifically, by following participating stroke survivors from the acute or subacute phase for an 18-month period, the project sought to:

1. assess the prevalence of mood and cognitive disorders prior to index stroke;

2. assess the accuracy and utility of brief cognitive tests and questionnaires addressing mood and subjective experiences post-stroke;

3. describe change in post-stroke cognitive function and mood over time.

APPLE was embedded within the NHS, with Ethics Committee and local Research & Development department approvals for all involved hospital sites (Research Ethics Committee number: 16/SS/0105). The project was funded by a joint grant from the Stroke Association and Chief Scientist Office of Scotland (funding reference: PPA 2015/01_CSO). The study protocol was registered on Research Registry (www.researchregistry.com; ID: researchregistry1018).

6.2 Participants and setting

Study participants were recruited from acute stroke units and outpatient stroke clinics of 11 hospital sites across the UK, including: the Glasgow Royal Infirmary (NHS Greater Glasgow & Clyde [NHS GGC]), Queen Elizabeth University Hospital (NHS GGC), Royal Alexandra Hospital (NHS GGC), University Hospital Monklands (NHS Lanarkshire), University Hospital Hairmyres (NHS Lanarkshire), Forth Valley Royal Hospital (NHS Forth Valley), Queen Margaret Hospital (NHS Fife), Perth Royal Infirmary (NHS Tayside), Aberdeen Royal Infirmary (NHS Grampian), Morriston Hospital (NHS Wales), and Charing Cross Hospital (Imperial College Healthcare NHS Trust). Participating sites admitted all adult cases of suspected stroke or TIA, regardless of premorbid physical and cognitive function.

Recruitment for APPLE took place between November 2016 and February 2019, involving both stroke survivors and informants. The Glasgow Royal Infirmary was the only site to be open throughout the whole recruitment period, with other sites opened at later stages. The last site to initiate recruitment was Charing Cross Hospital, beginning in January 2019.

Stroke participant selection criteria were intentionally broad, allowing to include any person over the age of 18 with a clinical diagnosis of stroke or TIA, provided they could converse in English prior to the event, they were not prisoners, and the responsible clinical team had no objection to their participation in the study. Stroke and TIA were defined as a focal, neurological event of presumed vascular cause, with the diagnosis made by a stroke specialist. All potential participants were also assessed for capacity to provide informed consent.

Patients who were eligible and willing to consider taking part in the study were given a Participant Information Sheet (PIS). After at least 24 hours, they would

be revisited by a member of the research team to discuss the study further, with an opportunity to ask questions and, if wishing to participate, would then sign a consent form. In addition to the core study, patients had the option to consent to future linkage of their data to clinical, electronic databases, and to provide blood and urine samples for biobanking (for NHS GGC sites only).

In cases where an eligible patient did not have the capacity to provide informed, written consent, consent was sought from a suitable proxy (a close relative or welfare guardian), while still involving the patient in the decision-making process as much as possible. Although the research team aimed to recruit patients within a week of their index stroke/TIA, following a protocol amendment, no specific cut-off timepoint was applied.

As well as direct assessments of stroke survivors, the APPLE study also involved collecting collateral information from suitable informants. Potential informants were indicated by patients, with an understanding that the former should know them and aspects of their day-to-day life well (e.g. spouse, child or carer). Similarly as in the case of patients, potential informants were given an appropriate PIS version, and if wishing to participate in APPLE – would sign a consent form on a separate occasion. According to the study protocol, recruitment of an informant for a patient was not conditional on the latter's participation in the study. In instances where an informant alone was consented, no data were collected through direct assessment of the patient, nor from the patient's medical records.

6.3 Assessments

The APPLE study involved 5 assessment timepoints – baseline, 1 month, 6 months, 12 months, and 18 months following recruitment. After the baseline assessment, subsequent follow-ups were scheduled within a two-week time window, either side of the initially intended completion date. For all visits, attendance was optional, meaning that participants could choose to skip any of the planned follow-ups, without withdrawing from the study overall.

Investigators from all sites received relevant instructional materials and training (face-to-face or remote, depending on the site location) for use of assessment measures. Baseline and 1-month assessments were conducted by research team

members from the site to which the stroke participant had been admitted. With the exclusion of participants from NHS Grampian and some from NHS Fife, later follow-up assessments (at 6, 12 and 18-months) were performed by our Glasgow team, including myself and three other researchers.

The assessment location was determined based on participant preferences. In the first instance, we would generally propose meeting at the local hospital site, where the assessment would take place in a private office. For this option, the stroke participant could make use of an arranged taxi service, and travel accompanied by the informant or another chosen person. In cases where travel did not suit the participant, following a risk assessment by the research team, we offered to arrange a home visit.

For the 6, 12 and 18-month follow-ups, a third alternative was to conduct the assessment over the telephone. This option was the only one available where a face-to-face follow-up was precluded by extensive travel distance. Towards the end of the APPLE study, with only 18-month follow-ups outstanding, conducting all assessments over the telephone became a necessity due to the COVID-19 pandemic.

All five stroke participant assessments involved a combination of cognitive tests and questionnaires regarding daily functioning, mood, and subjective experiences, while in the case of informants - only completion of questionnaires applied. The assessments were designed with an aim to ensure a thorough investigation of all aspects relevant to the study, while at the same time not overburdening participants. With the assumption that stroke participants' tolerance to length of assessments will improve alongside a gradual recovery after the acute phase of stroke, the baseline visit was made shortest, with the number of test and questionnaire items increasing over consecutive follow-ups until the 12-month assessment (the 18-month being identical).

6.3.1 Assessment materials

In relation to cognitive function, all stroke participant assessments incorporated a test very similar to the outcome measure described in Chapter 5 – a combination of items from the AMT-10 (364) and a short-form MoCA (77). The version used for the APPLE study differed by including one additional item

(counting down from 20), and in terms of scoring. Specifically, the question regarding date was scored out of three points instead of one (a separate point for each: day of the month, month and year), and the question regarding place was scored out of two points instead of one (one point for the exact location, e.g. name of hospital, street name and number of participant's home, and one for the name of the city). Consequently, the maximum total score for the test was 19 points. In the following text, I refer to this measure as the AMT-plus. Table 6-1 presents the test items and scoring system.

Table 6-1 Items and scoring for the AMT-Plus.

Item	Scoring
1. Age	1
2. Current time	1
3. Date: day, month, year	3
4. Place: location, city	2
5. Two-person recognition	1
6. Date of birth	1
7. Year World War I (or II) began	1
8. Current Prime Minister	1
9. Countdown from 20	1
10. Five-word delayed recall	1*
11. Clock draw: face, numbers, hands	3
12. Current news item	1
13. Months of the year in a reverse order	1
14. Verbal fluency (words beginning with "F")	1
	Sum: 19

*A point was assigned for the correct recall of at least four words.

To account for differences between assessments in inclusion of other measures, below I briefly describe the content of each assessment according to timepoint and version. A complete list of measures incorporated into each face-to-face assessment is presented in Table 6-2.

Table 6-2 Measures included in full-length and short stroke participant assessments.

	Assessments					
		Full-length versions			Short versions	
	Baseline	1-month	6-month	12/18-month	6-month	12/18-month
Objective measures of cognition						
Abbreviated Mental Test – plus (AMT-plus)	✓	✓	✓	✓	✓	✓
Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) (68)	✓	✓				
Oxford Cognitive Screen (OCS) (89)		✓				
Montreal Cognitive Assessment (MoCA) (77)					✓	✓
Animal Naming Test (393)			✓	✓		
Controlled Oral Word Association Task (COWAT) (394)			✓	✓		
Letter-Digit Substitution Test (LDST) (395)			✓	✓		
Consortium to Establish a Registry for Alzheimer's Disease (CERAD) Word List Memory Task, recall and recognition (396)			✓	✓		
Trail Making Test Parts A and B (397-399)			✓	✓		
CERAD Modified Boston Naming Test (396, 400)				✓		
Modified Rey-Osterrieth Complex Figure Test (401-403), copy and recall				✓		

Table 6-2 Measures included in full-length and short stroke participant assessments. *Continued*

	Full-length versions				Short versions	
	Baseline	1-month	6-month	12/18-month	6-month	12/18-month
Measures of function and activity						
Modified Rankin Scale (mRS) (171)	✓	✓	✓	✓	✓	✓
Barthel Index of Activities of Daily Living (404)	✓		✓			
Lawton Extended Activities of Daily Living	✓		✓			
"Frail Non-Disabled" (FiND) Instrument (405)	✓		✓	✓	✓	✓
Brief Fatigue Inventory (BFI) (406)		✓	✓	✓		
Brief Physical Activity Assessment (407)	✓			✓		✓
Stroke Impact Scale Short Form (SF-SIS) (408)				✓		
Measures of mood and subjective experience						
The Depression Intensity Scale Circles (DISCs) (409)	✓					
Patient Health Questionnaire - 2 (PHQ-2) (410)	✓	✓				
Generalised Anxiety Disorder Scale 2-item (GAD-2) (411)	✓	✓				
PHQ – Somatic, Anxiety and Depressive Symptom Scales (PHQ-SADS) (412-415)		✓				
Centre for Epidemiologic Studies Depression Scale Revised (CESD-R) (416)			✓	✓	✓	✓
EuroQol – 5 Dimension (EQ-5D) (417)			✓	✓	✓	✓
Medical Outcomes Study Social Support Survey (MOS-SSS), 4-item (418)	✓			✓		✓
Patient Reported Evaluation of Cognitive Status (PRECiS) (419)				✓		

6.3.1.1 Full-length stroke participant assessments

For the baseline visit, alongside AMT-plus the only other measure assessing for incident cognitive disorder was the CAM-ICU – a screening test for delirium (68). Included questionnaires, on the other hand, addressed the participants' pre-stroke state, encompassing brief measures of functional independence, frailty, anxiety and depression. At baseline, we additionally collected data on patient demographics, medical history, findings from acute clinical examinations, including stroke-related features, and laboratory test results. Appendix 8 presents the case report form (CRF) that we used to record most of this information.

Stroke subtype was defined using the Oxfordshire Community Stroke Project (OCSP) classification (420), based on a diagnosis made by the leading stroke physician. Stroke severity was determined based on the NIHSS (72); where the assessment had not been conducted by a clinician and recorded in the patient notes, the NIHSS was scored by an adequately trained member of the research team.

Unlike the baseline and all subsequent assessments, the 1-month visit included the Oxford Cognitive Screen (OCS; https://www.ocs-test.org/) as an additional measure of cognitive function (89). The OCS is a domain-specific tool, developed particularly to screen for cognitive impairment following stroke. As such, it was designed to be "aphasia and neglect-friendly", with an option to use a multiple-choice question format, and with test stimuli centred on a page. Domains assessed by the OCS include: language, attention, numerical skills, memory and praxis. In relation to questionnaires, measures included at 1 month referred to the participants' post-stroke condition, as was the case for all subsequent follow-ups. Compared to baseline, there were fewer scales related to activities of daily living, while additionally included were a measure of fatigue and a more comprehensive mood questionnaire.

At the 6-month follow-up, the OCS was replaced by a set of widely-used and validated domain-specific tasks, selected based on the National Institute for Neurological Disorders and Stroke (NINDS) and Canadian Stroke Network (CSN) recommendations for the assessment of vascular cognitive impairment (85). These task pertained to: semantic and phonemic verbal fluency, processing

speed, learning memory and executive functioning. Regarding questionnaires, the addressed aspects of the participants' condition and daily life were similar as for both the baseline and 1-month assessments, although in some instances different scales were used. Moreover, a quality of life measure was included. Further additions were made in the 12 and 18-month assessments, incorporating: word-retrieval and visuospatial ability and visuospatial memory tasks, as well as a stroke-specific measure of disability and health-related quality of life, and a questionnaire on the subjective experience of cognitive impairment and its impact.

6.3.1.2 Short versions of stroke participant assessments

Short versions were available as an alternative to 6, 12 and 18-month full-length assessments. They were administered at the request of the participant or at the researcher's discretion, in consideration of the participant's best interest, e.g. where completion of test and questionnaire items was associated with significant difficulty for the participant due to aphasia, fatigue, or poor general health. As presented in Table 6-2, short assessments included only a subset of measures used in the equivalent full-length version, while additionally comprising the MoCA. Tasks that overlapped between the AMT-plus and MoCA (e.g. clock draw, letter fluency) were assessed only once per visit, with an identical score recorded for both tests.

6.3.1.3 Telephone versions of stroke participant assessments

As described above, telephone versions were available as an alternative to 6, 12 and 18-month face-to-face visits. While the set of incorporated measures of function and mood matched that of the short versions, important adjustments were made for the assessment of cognition, with some AMT-plus items being omitted, and the inclusion of the modified Telephone Interview for Cognitive Status (TICS-M) (421, 422). In comparison to face-to-face follow-ups, the telephone assessments excluded the two-person recognition, clock draw, and five-word recall tasks, with the two former having no equivalent, and the latter being substituted by recall of a different, ten-word list.

6.3.1.4 Relevant amendments to stroke participant assessments

During preparation and design of my Biobank studies (Chapter 4), I proposed including measures of physical activity and social support as factors with plausible, yet understudied associations with post-stroke cognitive function. Wanting to avoid any significant increase in participant burden, after consultation with my Supervisor, we selected very brief scales. The questionnaires were added to the baseline, 12 and 18-month assessments for all versions in February 2018.

The included Brief Physical Activity Assessment is a two-item self-report measure, designed for use with adults in a primary healthcare setting (407). Through asking about the frequency of engaging in moderate and vigorous physical activity during a usual week, the tool's objective is to identify individuals who are insufficiently active in view of current recommendations. For assessment of social support, we used an abbreviated, 4-item version of The Medical Outcomes Study Social Support Survey (MOS-SSS) (418). The questionnaire addresses how often four types of social support are available to an individual, including: tangible support (material aid or assistance), positive social interaction (doing fun, enjoyable things with someone), emotional-informational support (emotional support and guidance or advice), and affectionate support (expression of love and affection).

6.3.1.5 Informant assessments

In many ways, informant assessments reflected those of the stroke participants. Administered at all timepoints apart from the 1-month follow-up, they could be posted, completed over the telephone, or face-to-face during the scheduled stroke participant visit. The assessments incorporated measures regarding changes in the stroke survivor's cognitive function, their functional status in view of activities of daily living, neuropsychiatric symptoms, mood, and quality of life. Questionnaires included in the baseline assessment referred to pre-stroke condition. Subsequent follow-ups additionally included one measure with questions about the informant, assessing caregiver burden. A full list of measures included at each informant assessment timepoint is presented in Table 6-3.

Table 6-3 Measures included in informant assessments.

	Assessments		
	Baseline	6-month	12/18-month
Questionnaire on Cognitive Decline in the Elderly (IQCODE) Short Form (423)	✓		✓
Aging and Dementia-8 (AD8) (424)	✓		✓
Modified Rankin Scale (mRS) (171)	✓	✓	✓
Barthel Index of Activities of Daily Living (404)	✓	✓	✓
Lawton Extended Activities of Daily Living	✓	✓	✓
"Frail Non-Disabled" (FiND) Instrument (405)	✓	✓	✓
Stroke Aphasic Depression Questionnaire for patients in hospital, 10-item (SADQ-H 10) (425)	✓		✓
Geriatric Depression Scale, 15 item (GDS-15) (426)	✓		✓
EuroQol – 5 Dimension (EQ-5D) (417)		✓	✓
Neuropsychiatric Inventory Questionnaire (NPI-Q) (427)			✓
The Zarit Burden Interview (428)		✓	✓

6.4 Study dropout

We recorded four reasons for study dropout: withdrawal of consent, loss to follow-up, death and "other", with the latter cause being specified by the researcher. Participant withdrawal mirrored the consent process. Specifically, participants who withdrew consent for future contact and follow-ups could (if applicable) additionally withdraw consent for data storage, linkage and access to medical records, and for storage of their samples in the biobank. The "other" option was typically selected in situations where we were unable to directly contact a participant, yet were aware of their current circumstances and felt that arranging future follow-ups would be inappropriate, e.g. end of life care situations. The same four dropout reasons were distinguished for informants. As a rule, if a participant dropped out of the study, we would withdraw their informant from APPLE as well. Such cases would be recorded as "other", with the exact cause documented.

6.5 Data processing

All CRFs were collated at the Academic Section of Geriatric Medicine in the Glasgow Royal Infirmary. Our team cross-checked CRFs completed by researchers from other sites and prepared all participant documents for transferral to the Robertson Centre for Biostatistics. The Centre, which designed the CRFs, was responsible for raising queries in case of suspected errors, data entry, and creating bespoke datasets, tailored to the requirements of our individual studies.

6.6 Recruitment outcome and follow-up completion

A total of 354 stroke participants and 151 informants were consented to APPLE, the majority being recruited from the Glasgow Royal Infirmary. The contribution of each site to stroke participant recruitment is presented in Figure 6-1.

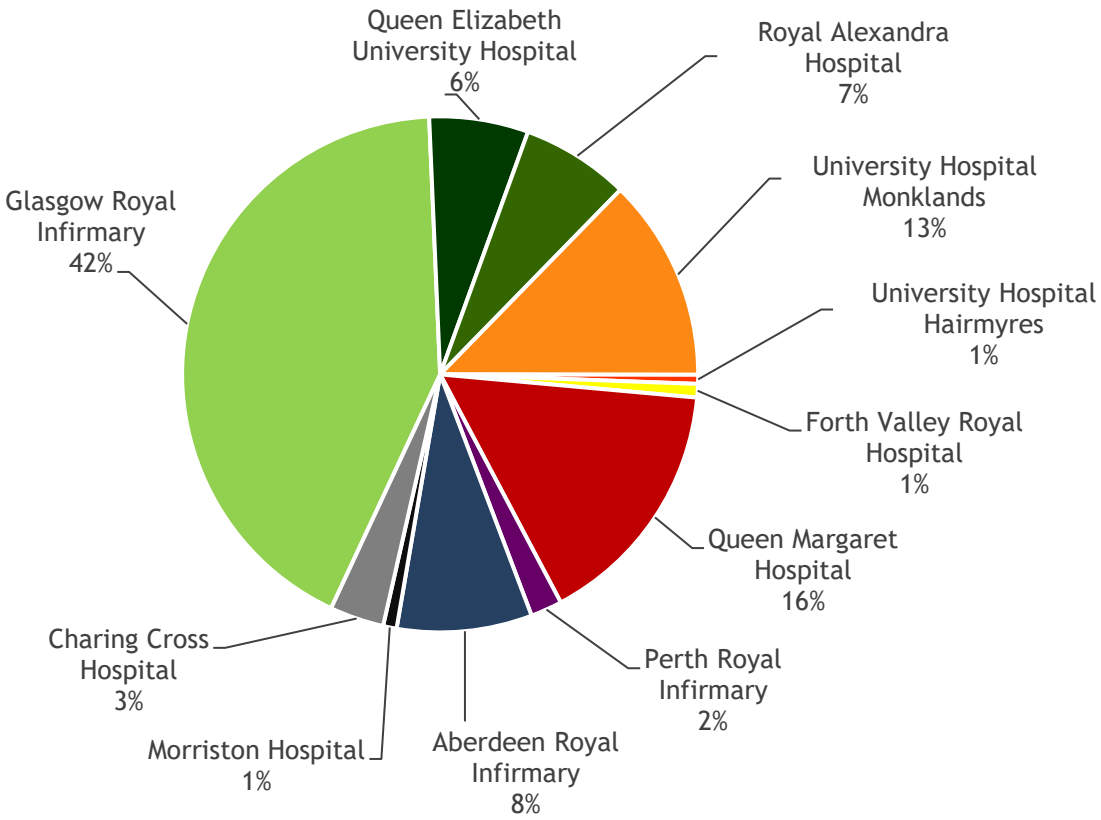


Figure 6-1 Percentage of stroke participants recruited to the APPLE study by each of 11 involved hospital sites.

Notes: NHS GGC sites are represented in green, NHS Lanarkshire in orange, NHS Forth Valley in yellow, NHS Fife in red, NHS Tayside in purple, NHS Grampian in blue, NHS Wales in black, and Imperial College Healthcare NHS Trust in grey.

6.6.1 Baseline study sample representativeness

I had access to patient screening logs from two participating sites – the Glasgow Royal Infirmary and the Royal Alexandra Hospital – to provide at least a partial indication of study sample representativeness. According to these records, during the period from November 2016 to February 2019, 352 patients were assessed for study eligibility at the Glasgow Royal Infirmary, of whom 149 (43%) consented to take part in APPLE (Figure 6-2). At the Royal Alexandra Hospital, open for recruitment between May 2018 and February 2019, 197 patients were assessed for eligibility, of whom 24 (12%) consented to study participation (Figure 6-3).

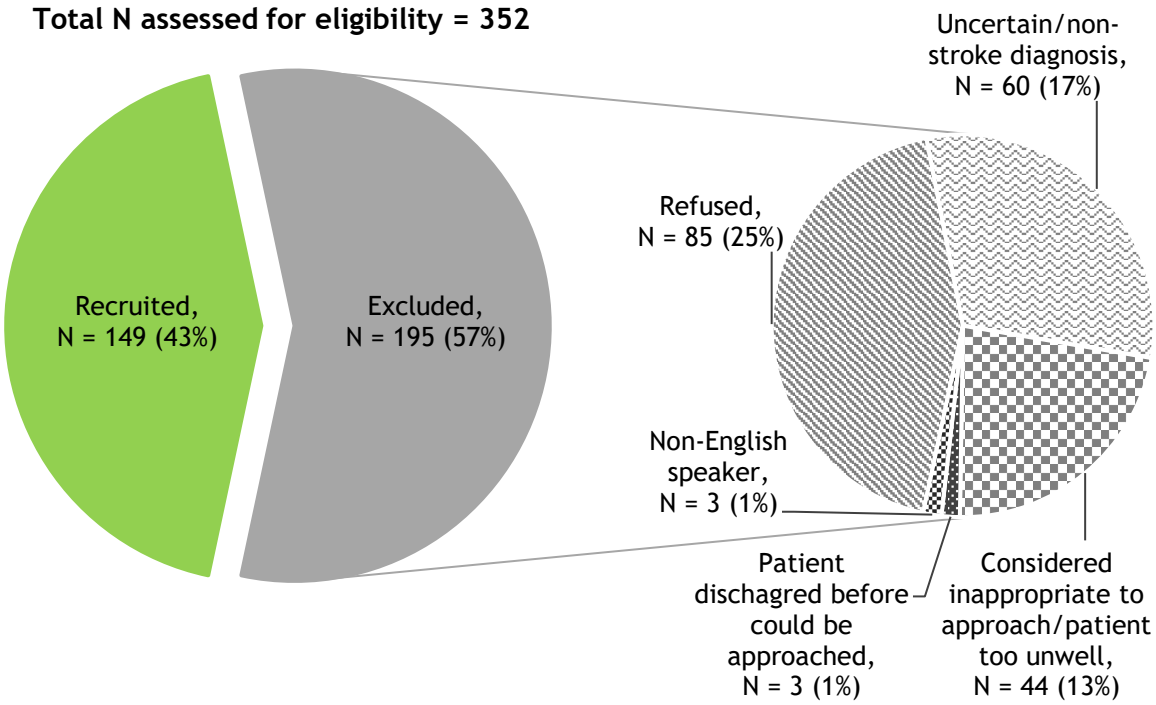


Figure 6-2 Patient screening and enrolment for the Glasgow Royal Infirmary.

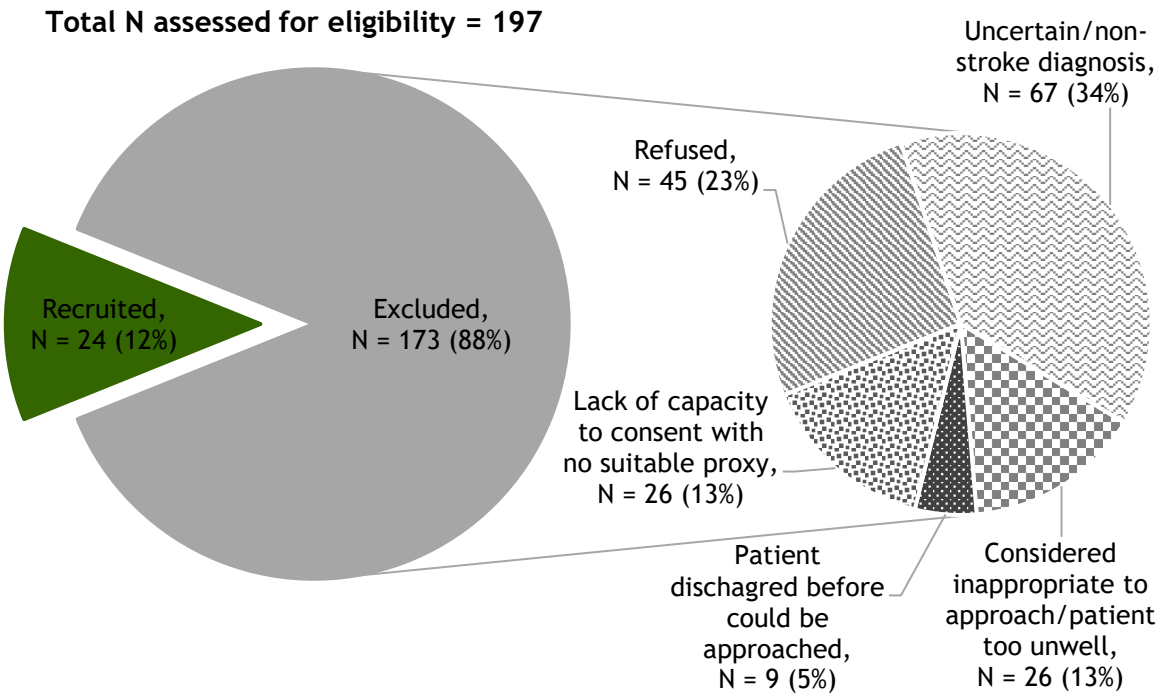


Figure 6-3 Patient screening and enrolment for the Royal Alexandra Hospital.

Table 6-4 includes selected characteristics of the APPLE baseline study sample. To allow further insight regarding participant representativeness, I present this information alongside data reported from the Sentinel Stroke National Audit Programme (SSNAP) (429), for the period between April 2017 and March 2018 (294). The clinical audit collects a minimum dataset for stroke patients admitted to every acute hospital in England, Wales, and Northern Ireland. Annually, information from approximately 85000 patients is submitted to the audit, representing over 90% of all stroke hospital admissions in the NHS (429). For the comparison of patient characteristics, it is important to note that SSNAP does not encompass cases of TIA.

Table 6-4 Characteristics of the APPLE baseline study sample presented for comparison with a national case mix of stroke patients, admitted to hospital between April 2017 and March 2018, as reported by the Sentinel Stroke National Audit Programme (SSNAP).

	APPLE cohort (N = 354)	SSNAP cohort (N = 83436)
Age (years)		
Median (IQR)	71.0 (59.8 – 79.3)	77.0 (67.0 – 85.0)
Missing	0	0
Sex		
Female, N (%)	157 (44.4%)	40764 (48.9%)
Missing	0	0
Diabetes, N (%)	86 (24.3%)	17860 (21.4%)
Missing	0	0
Hypertension, N (%)	191 (54.1)	45312 (54.3%)
Missing	1	0
Heart failure, N (%)	27 (7.6%)	4310 (5.2%)
Missing	0	0
Atrial fibrillation, N (%)	57 (16.1%)	16050 (19.2%)
Missing	0	0
Previous stroke/TIA, N (%)	87 (24.6%)	21602 (25.9%)
Missing	0	0
Pre-stroke mRS		
0, N (%)	177 (50.4%)	44328 (53.1%)
1, N (%)	56 (16.0%)	13264 (15.9%)
2, N (%)	61 (17.4%)	8992 (10.8%)
3, N (%)	52 (14.8%)	10058 (12.1%)
4, N (%)	5 (1.4%)	5248 (6.3%)
5, N (%)	0 (0.0%)	1546 (1.9%)
Missing	3	0
NIHSS score (stroke severity)		
Median (IQR)	2 (1 – 4)	5 (2 – 11)
Categories		
0, N (%)	84 (23.9%)	5288 (6.8%)
1 – 4, N (%)	183 (52.0%)	32624 (41.9%)
5 – 15, N (%)	75 (21.3%)	27428 (35.2%)
16 – 20, N (%)	6 (1.7%)	5633 (7.2%)
21 – 42, N (%)	4 (1.1%)	6925 (8.9%)
Missing	2	5538

IQR indicates interquartile range; mRS, modified Rankin Scale; NIHSS, National Institutes of Health Stroke Scale; TIA, transient ischaemic attack.

6.6.2 Follow-up completion

Focusing on participation rates at individual assessment timepoints, 269 (76.0%) stroke survivors took part in the 1-month follow-up, 220 (62.1%) in the 6-month follow-up, 185 (52.3%) in the 12-month follow-up, and 171 (48.3%) in the 18-month follow-up. We recorded that 158 stroke participants formally dropped out of the study, including 75 who withdrew consent (either directly or by proxy), 50 lost to follow-up, 24 who died, and 9 for whom the reason for dropout was "other" – in all cases related to a significant decline in health and/or end of life care.

6.7 Summary

APPLE was a multicentre, prospective cohort study, designed with an overall aim to improve our understanding of the neuropsychological sequalae of stroke. It involved a longitudinal 18-month follow-up of stroke survivors and their informants, with a focus on assessing cognition, mood, subjective experiences, and daily functioning. The work I conducted on APPLE constituted a central part of my doctoral training. Over the next two chapters, I describe how I used the data we collected to address the third of APPLE's work packages – describing change in post-stroke cognitive function.

Chapter 7 Trajectories of post-stroke cognitive change following stroke: A pilot study using the APPLE dataset. Part I: Rationale and methods

In the previous chapter, I described the design and assessment methods of a prospective inception cohort, involving longitudinal cognitive follow-up – the APPLE project. In this chapter, I present how I used the APPLE dataset to model the natural history of cognitive change following stroke. I devoted a separate part of my thesis to the methods of this study to allow a comprehensive and precise account of the extensive decision-making process that led to the final results. The methodological considerations I describe are important to the immediate interpretation of the study findings and, moreover, may serve to inform the design of future longitudinal research into post-stroke cognition.

7.1 Introduction

A common procedure employed in research aiming to identify predictors of post-stroke cognitive function relies on average-based estimates of associations between a factor of interest and a cognitive outcome, assessed at one specific timepoint. The uptake of this strategy was also evident across the prognostic model development studies I described in my systematic review (Chapter 3). The core characteristic of this approach is that it is “variable-centred”, as the focus is on the relationships among variables (430). Alongside practical advantages and relative ease in interpretation of results, there is, however, an important limitation to this popular strategy when applied to post-stroke cognition - it does not reflect the heterogeneity in the process of cognitive change.

Clinical observations and research findings suggest a dynamic and varied pattern of cognitive change following stroke. Although many stroke survivors will experience an initial period of cognitive recovery, only some will continue to improve in the longer-term, while for others this process will reach a plateau or even shift towards a declining trend (431-434). This suggests heterogeneity at both a between- and within-individual level. Regarding the former, even where at a specific timepoint the same outcome has been apparently achieved by two individuals, the path to it may have differed.

For example, as illustrated in Figure 7-1, a person may be diagnosed with cognitive impairment one year after a stroke either as a result of declining from a previous state of having no cognitive impairment (Person A), or following a significant but incomplete recovery from severe initial deficits (Person B). Despite being categorised into the same outcome group, arguably individuals representing these different patterns of change are also likely to differ in terms of characteristics relevant to post-stroke cognitive function, and possibly in how their condition will progress.

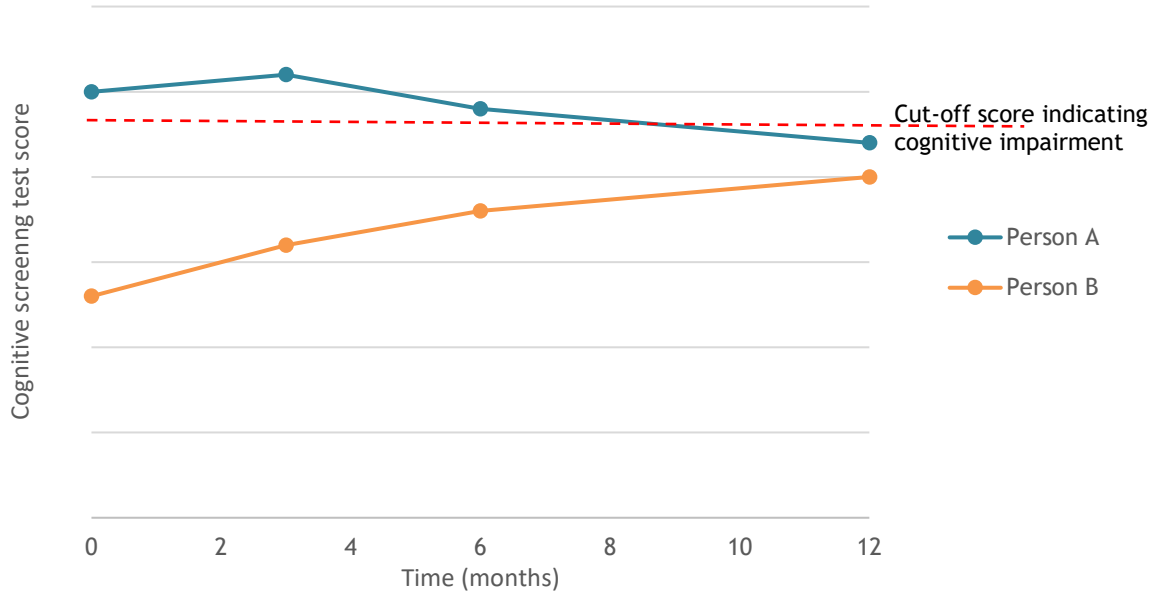


Figure 7-1 Hypothetical trajectories of cognitive change for two individuals classed as having cognitive impairment 12 months following stroke.

Concurrently, on a within-individual level, cognitive change is unlikely to follow a constant trajectory over time, with potential for variability in rate and, in some cases, even shift in direction. Indeed, observations from longitudinal follow-up studies in stroke cohorts indicated that for some participants the cognitive status identified (impaired or intact) switched up to two times within a three-year period (435, 436). In a similar study, the authors even reported instances of reversion from a vascular dementia diagnosis to milder forms of impairment, and for one person - to unimpaired cognition (391). These observations illustrate what additional insights can be gained through investigating individual trajectories of cognitive change.

However, despite the advantages of the approach used in these studies, it still does not allow to capture the full complexity of the subject matter. As a consequence of defining the outcome as binary (cognitive impairment vs no cognitive impairment) (435, 436), any changes that do not result in crossing a diagnostic threshold are missed. A second, partially connected issue relates to applying a priori simplifications to classifying participants based on cognitive change. This was reflected in basing the decision about categorisation solely on change over two chosen timepoints (391, 435, 436). As a result, two individuals may be grouped together, for example, as “improvers”, despite presenting meaningful differences in severity of initial cognitive impairment, level of residual difficulties at follow-up, or rate and specific pattern of change in between those timepoints (Figure 7-2).

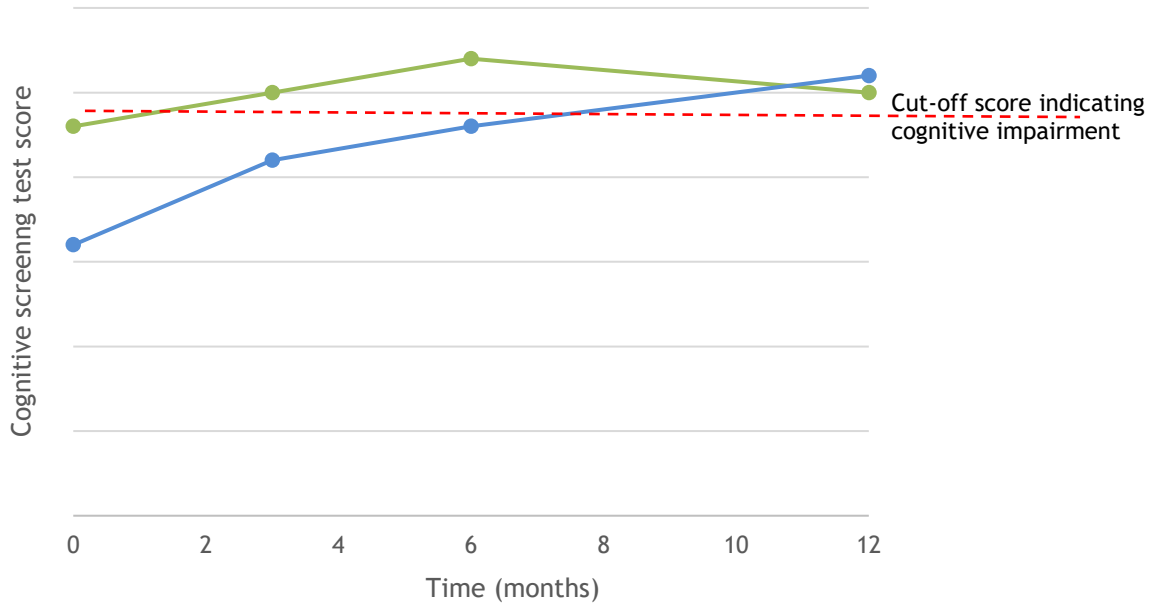


Figure 7-2 Hypothetical trajectories of cognitive change following stroke for two individuals classed as “improvers” (change in status from impaired to intact cognition) through a comparison of baseline and 12-month cognitive test scores.

Such challenges, present in many areas of research focusing on processes and change in condition over time, have contributed to increasing interest in integrating variable-centred analyses with a person-centred approach (430). This goal can be achieved through use of novel analytical strategies, for example latent growth modelling techniques, such as growth mixture modelling (GMM) (437, 438) and latent class growth analysis (LCGA) (439). Based on recognition of a population’s heterogeneity, these methods are applied to identify distinct,

homogenous subpopulations (classes) within it. This allows interindividual differences in intraindividual change over time to be captured (440). In addition to visualising the unique features of distinguished trajectory shapes, it is then also possible to determine what factors are uniquely associated with exhibiting a particular pattern of change.

The aim of this study was to explore the application of a latent growth modelling approach to identify meaningful trajectories of post-stroke cognitive change in the APPLE dataset. At present, data entry, cleaning, and quality control are still ongoing, with information we had collected at the 18-months assessments not yet released by the clinical trials unit (the Robertson Centre for Biostatics) for APPLE. Therefore, the results of the following analyses, involving data from baseline to the 12-month follow-up, will serve to inform the design and choice of specific statistical solutions for a subsequent, full-scale study, utilising information from all five APPLE assessment timepoints.

7.2 Methods

The conduct of the present study compromised of four main sections: i) a factor analysis to derive a latent cognitive variable based on raw cognitive scores, to serve as the outcome; ii) the selection and implementation of an approach for handling missing outcome data; iii) the application of a latent growth modelling technique to identify distinct trajectories of cognitive change over time; and iv) the investigation of predictors of trajectory class membership. Below I provide a detailed description of the rationale and steps involved in each study section, with special consideration to Guidelines for Reporting on Latent Trajectory Studies (GRoLTS) (441). I begin by presenting my choice of a latent growth modelling technique - although this issue specifically pertains to a later stage of my study, it was a primary decision, affecting methodological considerations from the outset.

7.2.1 Choice of modelling approach

Given the pilot nature of the study, I chose to perform a LCGA. This approach is recommended as a first step prior to attempting the more complex GMM (440). Through utilising SEM techniques, the goal of both methods is to identify trajectory-based classes within a sample, where individuals within one class are

most similar to one another, while at the same time being most different from individuals in all other classes (442). However, the difference is that LCGA entails an assumption that growth trajectories (here, counterintuitively signifying any pattern, not only an increasing function) are homogenous within a class, while GMM allows to account for within-class variation.

The greater flexibility offered by GMM comes at a cost of increased computational demand, more frequent convergence issues (i.e. inability to generate an admissible model solution), and need of greater statistical power. Nonetheless, this approach may result in a more accurate reflection of real-world patterns of change where within-class heterogeneity exists, which in most contexts cannot be ruled out. Therefore, I intend to test the use of GMM in the full-scale study, for which the inclusion of an additional assessment timepoint would translate to an increase in power.

7.2.2 Participants

The study sample comprised APPLE participants. I applied no selection criteria based on individual characteristics. However, I considered that to develop an accurate model using LCGA it is key that assessment timepoints are similar across participants, relative to an objective, study-independent starting point (443). Given the design of APPLE, described in the previous chapter, a particular concern was that, relative to index stroke, the baseline assessment of some participants would be closer in time to the 1-month assessment of the majority.

Conversely, introducing strict selection criteria based on recruitment time could have led to increased sample bias, as plausibly informing stroke survivors and their relatives about APPLE may have been delayed where the patient was very unwell in the acute phase. Therefore, I decided to apply a cut-off for inclusion of baseline data collected four weeks post-stroke or less. As a result, I included 343 stroke participants in the present study, equivalent to 96.9% of the original APPLE sample.

As indicated in Chapter 5, there are no consensus guidelines regarding sample size requirements for statistical approaches based on SEM. Multiple factors affect this issue, including the study design, number of parameters to be estimated, distribution of variables, missing data, and effect sizes of tested

associations (444). However, based on previous research, it seems that for a dataset including four assessment timepoints, a sample of 200 participants is sufficient for generating unbiased model estimates using LCGA (443, 445, 446).

The appropriateness of the sample size for investigating predictors of class membership was a separate issue, which could not be addressed before determining the number of distinct trajectories. Overall, in view of the exploratory objective of this research, I decided to attempt all planned analyses, regardless of the associated statistical power. However, where appropriate, I would introduce adjustments to compensate for sample size limitations.

7.2.3 Cognitive data

I used the AMT-plus as a measure of cognitive outcome. Although scores from the neuropsychological battery used in APPLE would have offered greater insight into the participants' cognitive function, the AMT-plus was the only test repeated across all face-to-face assessments, satisfying an essential requirement for LCGA. Moreover, as described in Chapter 6, most of original AMT-plus items (14 of 19) were included in the telephone versions of 6, 12, and 18-month assessments. Here, the lack of data on unincluded items could have been plausibly assumed as missing completely at random (no systematic differences between missing and observed values) or missing at random (where any systematic differences can be explained by observed data, e.g. if a telephone assessment was opted for due to the participant's functional status) (296). Therefore, I considered that in this case the use of an imputation procedure would be appropriate.

This was one of two main arguments for computing factor scores rather than using AMT-plus sum scores for the LCGA. A factor score can be estimated for a participant with missing values on specific test items based on observations available in the dataset. The second argument related to weighting of items. For sum scores, items are assigned equal weight, that is, they are assumed to contribute an equal amount of information to the measured construct (447). However, from a clinical perspective, this seems to be unlikely for the AMT-plus. For example, not knowing the city one is in or the current month is likely more indicative of a cognitive disorder than not knowing the exact date (448).

I describe the process of obtaining factor scores to represent cognitive function in the section below. The dataset I used contained information on all 19 AMT-plus items, coded as separate, binary variables (correct vs incorrect). I included every assessment type – full-length, short and telephone versions. Similarly as in my study described in Chapter 5, in cases where a participant could not respond to a test item due to an existing impairment (e.g. aphasia or limb weakness), I assigned a score of 0 for that item.

7.2.4 Transforming raw cognitive scores into the outcome variable: a factor analysis approach

7.2.4.1 Factor analysis procedure

An inspection of relationships between the 19 cognitive test items revealed extremely high correlations (above 0.9) and instances where no participant presented a specific combination of responses for a pair of variables. For example, at the 1-month assessment, there were no cases where providing an incorrect date of birth cooccurred with correct responses regarding age, time, date, or place. Consequently, it was not possible to produce factor scores based on a model simultaneously including all items.

Aiming to find a solution that retained as much information as possible, I followed the iterative process described below. After each step (with the exclusion of the first one), I verified whether an admissible model solution could be found, or whether further adjustments were necessary.

1. Combining all items from the clock draw task, creating a new, four-level item (range: 0 to 3).
2. Combining items that could be justifiably grouped together based on both high intercorrelations and clinical interpretation:
    a) Eight items used in the assessment of consciousness (e.g. GCS and NIHSS) and/or relevant to orientation – age, date of birth, time, date, month, year, place and city (range: 0 to 8);
    b) Two items relevant to assessing attention – counting down from 20 and listing months of the year backwards (range: 0 to 2).

3. Collapsing orientation and attention scores to create binary items, with cut-offs based on sample distributions;

   a) For orientation – one assigned for sum scores of seven and eight, zero for scores below seven; as a result, across the four assessment timepoints, the percentage of participants who were assigned a point ranged from 81.8% to 90.9%;

   b) For attention – one assigned for a sum score of two, zero for scores below two; the percentage of participants who were assigned a point ranged from 77.7% to 84.9%.

4. Removing item on recognition of two people.
5. Removing item on naming current prime minister.

The last two steps were necessary as the high correlations between these items and the orientation item, pertaining to at least one assessment timepoint, precluded from their simultaneous inclusion in a single model. Consequently, the final model structure incorporated seven items: orientation, attention, recent news item, five-word recall, year WWI began, clock draw, and verbal fluency.

7.2.4.2 Assessing properties of the derived cognitive latent variable

As all test items (indicators) were categorical, I used a robust weighted least squares estimator (WLSMV). With this estimator, missing data are handled based on a pairwise present analysis, where all available observations are used to estimate correlations between each pair of items (367). I assessed model fit based on the chi-square statistic (good fit indicated by non-significance), Root Mean Square Error of Approximation (RMSEA; good fit: ≤ 0.06), Comparative Fit Index (CFI; good fit: ≥ 0.95), and Tucker-Lewis Index (TLI; good fit: ≥ 0.95).

Given the longitudinal nature of the data, it was further necessary to examine measurement invariance. This is to ensure that the relationship between test items and the latent construct that underlies them (here, cognitive function) remains unchanged across timepoints or, in other words, that the meaning of the investigated construct is the same for each assessment occasion (449, 450). Without satisfying this condition, changes in factor scores over time cannot be reliably attributed to actual changes in the construct.

For categorical indicators, assessing measurement invariance involves comparing a configural model to a scalar model (366). For the configural model, I specified the same latent factor structure (involving the same seven items) for each timepoint, while allowing for loadings and thresholds of indicators to be freely estimated. For the scalar model, I constrained loadings and thresholds to be equal for corresponding items across timepoints. Invariance is considered to be achieved if the chi-square statistic for the scalar model is not significantly worse than for the configural model, relative to the change in degrees of freedom. I conducted the factor analysis using Mplus version 8.3. The code I developed for both models is presented in Appendix 9.

7.2.4.3 Factor analysis results

For the configural model, including seven cognitive test items, considered indices suggested good model fit: $X^2 = 325.91$, p = 0.16; RMSEA = 0.02; CFI = 0.99; TLI = 0.99. After applying additional equality constraints for the scalar model, model fit did not significantly deteriorate (chi-square difference: 8.5, degrees of freedom: 15, p = 0.90), indicating that measurement invariance had been achieved. Table 7-1 presents estimated factor loadings in a descending order, while Table 7-2 includes descriptive statistics for computed factor scores.

Table 7-1 Unstandardised loadings for the cognitive function latent construct.

Item	Unstandardised loading
Attention	1.01
Orientation	1.00
News item	0.74
Word recall	0.65
Clock draw	0.60
Verbal fluency	0.55
World War I	0.47

Table 7-2 Descriptive statistics for estimated factor scores across assessment timepoints.

	Baseline	1-month	6-month	12-month
Range	−2.55 to 1.60	−2.36 to 1.92	−2.03 to 1.54	−1.39 to 1.34
Mean (SD)	−0.08 (0.78)	0.22 (0.84)	0.25 (0.64)	0.40 (0.47)
Missing	2	81	128	162

SD indicates standard deviation.

7.2.5 Missing outcome data due to study dropout

7.2.5.1 Considering strategies for handling missing outcome data

For the purpose of this study, I defined “dropout” as any case, where once an assessment had been missed, no subsequent follow-up was completed. This included instances of participants repeatedly requesting to skip visits, without formally withdrawing from the study.

In latent growth modelling research, the most popular approach to handling missing data due to dropout is full information maximum likelihood estimation (FIML) (441). This method involves parameter estimation using all the data that are available, without imputing specific missing values (451, 452). A key advantage of applying FIML is that each participant, regardless of whether they completed the study, is assigned to a latent class. However, unbiased estimates can only be produced where data are missing completely at random or at random.

Missingness at random tends to be assumed on the basis that lack of follow-up data for a participant is conditional on outcome information collected at previous, completed assessment(s) (441). However, it is becoming increasingly recognised that at least for certain outcomes and populations, this assumption is unlikely to hold (453-455). Dropout may indeed relate to deterioration from previous status, which would not be accurately reflected using FIML.

This concern seems highly relevant to investigating post-stroke outcomes. It seems that what has the greatest possibility of being captured in participant data, is the initial improvement in function, typically seen in the acute and subacute phase of stroke. If after that stage it is mainly stroke survivors with better and/or improving cognitive function who remain in the study, estimated trajectories may present an overoptimistic view of post-stroke cognitive change over time, as compared to patterns occurring within a real-world, unselected stroke population.

There is no direct method to test whether data are missing at random or not. However, comparing characteristics of participants who remained in a study to those who dropped out may clarify whether the latter were indeed at greater

risk of outcome deterioration. The results of such a comparison may also depend on the reason for dropout (456). In view of this, I utilised available APPLE records to distinguish between cases lost to follow-up due to death or end of life care, and cases lost for other reasons.

I compared the three participant groups (study completers and two groups with lost to follow-up) based on factors considered as relevant to cognitive decline. Regarding demographics, I accounted for age, sex, and education (in years). For health-related factors, I included: BMI (kg/m^2), medical conditions (diabetes mellitus, vascular disease, heart failure, atrial fibrillation, previous stroke or TIA, renal disease, prior cognitive impairment, history of mood disorders, and history of substance abuse, including alcohol and illicit drug use), pre-stroke functional dependency as indicated by a mRS score of above two (171), and lifestyle factors (smoking status and self-reported physical activity). In terms of pre-stroke status, I additionally considered the subjective level of received social support (based on the MOS-SSS questionnaire). Finally, accounting for acute presentation, I included stroke severity as measured by the NIHSS (71, 72), and baseline cognitive function (both raw AMT-plus sum scores and factor scores).

To test for potential differences, I used chi-squared and Fisher's exact tests for categorical variables, and the Mann-Whitney U test for continuous ones. I noted whether group differences were significant at $p < 0.05$, as well as after accounting for multiple comparisons, using the Holm-Bonferroni technique (Holm, 1979). Results indicating that dropout was associated with risk factors for cognitive decline would support complementing a latent trajectory analysis using FIML with an alternative approach to handling missing data, to reflect poorer outcome in lost cases.

Given the pilot nature of this study, I intended to adapt a computationally nondemanding method, described in a publication on trajectories of functional limitations in later life (453). There, for participants who died during the study follow-up, the authors assigned (with random noise) functional limitations greater by one standard deviation (1 SD) from the timepoint-specific sample mean, for every missed assessment until death.

For the APPLE sample, I assumed it was likely for both dropout groups to appear at higher risk of cognitive decline, however, with the difference from study-completers being greater for participants who died/were in terminal care. If my assumption were confirmed, I planned to assign a cognitive function score lower by 2 SD from the timepoint-specific sample mean for dropout due to death/terminal care, and lower by 1 SD for dropout due to other reasons.

7.2.5.2 Results of group comparisons according to dropout status

Descriptive statistics for the total study sample and participant groups distinguished based on study dropout are presented in Table 7-3. Results of the univariable analyses indicated that both dropout groups differed from participants who completed APPLE assessments, even after adjusting for multiple comparisons. However, there was little overlap between groups regarding variables to which these differences applied to.

Summarising results significant at least at $p < 0.05$, participants who dropped out due to death or end of life care were on average older than study-completers, were more frequently diagnosed with heart failure, atrial fibrillation and previous stroke/TIA, less physically active, and less likely to be functionally independent. Although the latter also applied to the group with dropout due to other reasons, remaining differences included a higher proportion of female participants, lower average education, more cases of prior cognitive impairment and mood disorders, and poorer baseline cognitive function.

Table 7-3 Descriptive statistics for study sample and group comparison by dropout status.

	Total sample (N = 343)	Group comparison by dropout status		
		Study-completers (N = 199)	Dropout due to death/terminal care (N = 25)	Dropout due to other reasons (N = 119)
Age (years), Mean (SD)	69.2 (12.8)	68.1 (12.8)	77.0 (8.2)**	69.5 (13.2)
Missing	0	0	0	0
Sex				
Female, N (%)	153/343 (44.6%)	76/199 (38.2%)	9/25 (36.0%)	68/119 (57.1%)**
Education (years), Mean (SD)	12.0 (3.4)	12.5 (3.7)	12.0 (2.0)	11.3 (2.8)**
Missing	29	12	1	16
BMI				
<18.5 (underweight), N (%)	9/335 (2.7%)	2/197 (1.0%)	1/25 (4.0%)	6/113 (5.3%)
18.5 to 24.9 (normal), N (%)	106/335 (31.6%)	62/197 (31.5%)	8/25 (32.0%)	36/113 (31.9%)
25.0 to 29.9 (overweight), N (%)	120/335 (35.8%)	71/197 (36.0%)	10/25 (40.0%)	39/113 (34.5%)
>30 (obese), N (%)	100/335 (29.9%)	62/197 (31.5%)	6/25 (24.0%)	32/113 (28.3%)
Smoking status				
Never, N (%)	127/341 (37.2%)	75/199 (37.7%)	10/24 (41.7%)	42/118 (35.6%)
Former, N (%)	140/341 (41.1%)	88/199 (44.2%)	10/24 (41.7%)	42/118 (35.6%)
Current, N (%)	74/341 (21.7%)	36/199 (18.1%)	4/24 (16.6%)	34/118 (28.8%)
History of substance abuse, N (%)	37/341 (10.9%)	20/198 (10.1%)	1/25 (4.0%)	16/118 (13.6%)
Diabetes, N (%)	85/343 (24.8%)	46/199 (23.1%)	8/25 (32.0%)	31/119 (26.1%)

Table 7-3 Descriptive statistics for study sample and group comparison by dropout status. *Continued*

	Total sample (N = 343)	Group comparison by dropout status		
		Study-completers (N = 199)	Dropout due to death/terminal care (N = 25)	Dropout due to other reasons (N = 119)
Hypertension, N (%)	183/342 (53.5%)	102/199 (51.3%)	14/25 (56.0%)	67/118 (56.8%)
Vascular disease, N (%)	98/343 (28.6%)	56/199 (28.1%)	11/25 (44.0%)	31/119 (26.1%)
Heart failure, N (%)	26/343 (7.6%)	9/199 (4.5%)	7/25 (28.0%)**	10/119 (8.4%)
Atrial fibrillation, N (%)	56/343 (16.3%)	25/199 (12.6%)	9/25 (36.0%)**	22/119 (18.5%)
Previous stroke/TIA, N (%)	85/343 (24.8%)	41/199 (20.6%)	12/25 (48.0%)**	32/119 (26.9%)
Renal disease, N (%)	41/343 (12.0%)	23/199 (11.6%)	6/25 (24.0%)	12/119 (10.1%)
Prior cognitive impairment, N (%)	26/343 (7.6%)	10/199 (5.0%)	1/25 (4.0%)	15/119 (12.6%)*
History of mood disorders, N (%)	91/343 (26.5%)	44/199 (22.1%)	8/25 (32.0%)	39/119 (32.8%)*
Pre-stroke mRS				
Dependency (mRS > 2), N (%)	56/340 (16.5%)	23/197 (11.7%)	9/25 (36.0%)**	24/118 (20.3%)*
Physical activity (range: 0 to 8)				
Mean (SD)	2.3 (2.6)	2.7 (2.8)	0.8 (1.1)*	2.1 (2.3)
Missing	178	102	13	63
Social support (range: 4 to 20)				
Mean (SD)	17.2 (3.8)	17.8 (2.9)	15.1 (5.1)	16.5 (4.5)
Missing	178	102	13	63

Table 7-3 Descriptive statistics for study sample and group comparison by dropout status. *Continued*

	Total sample (N = 343)	Group comparison by dropout status		
		Study-completers (N = 199)	Dropout due to death/terminal care (N = 25)	Dropout due to other reasons (N = 119)
Stroke severity (NIHSS; range: 0 to 42)				
Mean (SD)	3.3 (4.3)	2.9 (3.9)	3.6 (3.4)	3.9 (5.0)
Missing	2	1	0	1
Categories				
No stroke signs, N (%)	82/341 (24.0%)	53/198 (26.8%)	4/25 (16.0%)	25/118 (21.2%)
Mild, N (%)	179/341 (52.5%)	106/198 (53.5%)	13/25 (52.0%)	60/118 (50.8%)
Moderate, N (%)	70/341 (20.5%)	35/198 (17.7%)	8/25 (32.0%)	27/118 (22.9%)
Severe, N (%)	10/341 (2.9%)	4/198 (2.0%)	0/25 (0.0%)	6/118 (5.1%)
Baseline cognition				
Sum score (range: 0 – 19), Mean (SD)	15.0 (3.6)	15.6 (3.2)	14.6 (3.5)	14.2 (4.1)**
Missing	14	6	0	8
Factor score, Mean (SD)	-0.1 (0.8)	0.1 (0.7)	-0.2 (0.8)	-0.3 (0.8)**
Missing	2	1	0	1

Note: Univariable comparisons were made between the two dropout groups and the group of study-completers.

*significant at $p < 0.05$; **significant after applying Holm-Bonferroni correction for multiple comparisons

BMI indicates body mass index; mRS, modified Rankin Scale; NIHSS, National Institutes of Health Stroke Scale; SD, standard deviation; TIA, transient ischaemic attack.

7.2.5.3 Finalising the approach for handling missingness

I had initially considered assigning a lower cognitive score for drop out due to death (or end of life care) compared to drop out due to other reasons. However, based on my findings, an assumption that one dropout group was at greater risk of cognitive decline than the other seemed arbitrary. Moreover, it was possible that not all participants had been correctly classified according to dropout status. Particularly, where participants withdrew consent or were lost to follow-up at earlier timepoints, deaths occurring later during the study period may not have been recorded. For these reasons, I decided to substitute missing outcome values to represent a similar level of cognitive function for both groups. Initially, this was around two standard deviations below the sample mean (−2 SD) for a specific assessment timepoint. To reduce the artificiality of a dataset generated in this way, I added a component of random gaussian noise. Consequently, participants with missing outcome data due to dropout would be randomly assigned a score from approximately −2.5 SD to −1.5 SD, with most values being close to −2 SD, and fewest at either extreme.

However, an attempt to apply this method proved that for many participants who dropped out, it would fail to reflect a realistic decline in cognitive function. This was due to the relatively high variability in cognitive function factor scores. As such, value substitution for participants with high cognitive function, as measured at previous, completed assessments, would indicate a drastic decline, while for participants with poor cognitive function – an improvement. I therefore opted to divide the participants according to baseline factor scores, forming three groups: with scores within ±1 SD (N = 232), with scores below −1 SD (N = 55), and with scores above 1 SD (N = 56). I then generated values to use for substitution based on the timepoint-specific mean of the group the subject had been assigned to.

A final consideration was which missed assessment values should be substituted. From a clinical perspective, it seemed that replacing values for each timepoint from dropout to either death or the end of follow-up would lead to implausible patterns of cognitive change over time. This was partially due to decreasing dispersion of values across time, meaning that with each timepoint, a subtraction of around −2 SD resulted in a smaller difference from the group mean. To illustrate the implications of this with an example, if a participant

withdrew from the study after baseline, and every subsequent missing outcome value were substituted, their cognitive function would be found to decline at 1-month as compared to baseline, reaching its lowest point, after which it would consistently improve.

At the same time, given the exploratory nature of this study, it seemed uncertain what patterns post-stroke cognitive change would be likely to follow, and whether more than one pattern occurred. I therefore decided to apply an approach where missing value substitution could not dictate the whole shape of a participant's trajectory, yet rather would steer it towards a direction indicating cognitive decline. Specifically, for participants who according to APPLE records were still alive at the 12-month assessment, I only substituted missing values for this last follow-up, while for participants who died within the duration of the study – I substituted the missing value for the timepoint at which their death had been noted. All remaining missingness was handled based on FIML.

7.2.6 Main analyses

The main analyses comprised of modelling trajectory classes of cognitive change over time and identifying predictors of class membership. I carried out this part of my research applying a standard three-step method, as specified in the GRoLTS publication (441):

1. I determined the number of latent trajectory classes without potential predictors of class membership;

2. I saved the most likely class membership for each participant as a new variable, adding it to the original dataset;

3. I investigated predictors of class membership in separate analyses, involving logistic regression and mediation models.

7.2.6.1 Identification of latent trajectory classes

In this key part of the study, I firstly gave special consideration to differences in timing of assessments across participants. Despite excluding subjects who were recruited to APPLE more than four weeks following stroke, there was still non-negligible variability in when participants completed study visits. This seemed particularly relevant for early assessment, occurring over a period when changes in cognitive function would likely be most dynamic.

For example, the average length of time from index stroke to baseline assessment was approximately one week, and ranged from one to twenty-eight days, with a standard deviation of five days. For the 1-month follow-up, the interval from stroke onset ranged from 24 to 69 days (M = 41, SD = 8). If unaccounted for, individual differences in assessment timepoints are likely to lead to misestimation of model parameters, and may hinder successful model convergence (443). Therefore, I conducted a time-unstructured analysis, where individual assessment timepoints are recorded for each participant and included as variables in the model (here, an additional four variables) (457).

A second central issue related to the functional form of trajectories, capturing cognitive change over time. In studies using LCGA, the most commonly implemented approaches are polynomial functions – linear (straight line), quadratic (one curve), cubic (two curves), and incorporating more curves. The type of growth function that can be modelled is influenced by the number of assessment timepoints (444). With four timepoints, it was possible to specify either a linear or quadratic pattern of change. In selecting the optimal solution, I followed a similar approach as for deciding on the final number of latent trajectory classes.

I investigated models including from one to six trajectories, assuming that a higher number would pose a challenge in view of the available sample size. In line with current recommendations, I compared the models on multiple aspects. Firstly, I assessed model fit indices: the Akaike information criterion (AIC) (458), the Bayesian information criterion (BIC) (459), and the sample size-adjusted BIC (SSA-BIC) (460). For these statistics, a lower value indicates better model fit. I then considered classification accuracy. Entropy is a summary measure reflecting how well classes are separated from one another, and the confidence

with which participants are assigned to a specific class (441, 461, 462). Estimates range from 0.00 to 1.00; values approaching 1.00 are preferred, with a widely applied acceptability threshold of 0.80 (463). Classification accuracy is also estimated for individual classes, according to model-based (posterior) probabilities (440). I presented this measure as a percentage, where values approaching 100% are favoured. It is important to note that for every participant a probability is estimated for belonging to each identified class – the class with the highest probability is the one the person will be assigned to.

Another aspect relevant to model comparison includes the number of individuals assigned to a particular latent trajectory class. Generally, solutions where less than 1% of the study sample represents any given class are rejected in favour of a model with fewer trajectories (440). A similar approach is taken where identified trajectories highly resemble one another in shape, and any distinguishing features are difficult to identify. This also relates to the final assessment component, focusing on clinical plausibility and interpretation.

An example of a Mplus code I developed for a three-class solution is provided in Appendix 10. An important feature of the code is the specified number of random sets of starting values and the number of final optimisations (two values for STARTS), i.e. iterations based on maximum likelihood parameter estimation, which I increased from default settings (specifically, from STARTS = 10 2 to STARTS = 100 10) (464). This allowed for a more thorough investigation of multiple solutions, and in turn increased the probability of obtaining an optimal one rather than a solution based on local maxima. Ideally, the iteration is to result from successful convergence on the global maximum solution (440). Once I selected the most favourable model from considered alternatives, I verified this condition by testing whether the parameter estimates would be replicated for the two best obtained loglikelihood values (440, 444).

7.2.6.2 Prediction of class membership

This part of the study involved a logistic regression analysis, with latent trajectory class membership constituting the outcome. As follows, the type of developed model – binary or multinomial – depended on the number of classes in the final solution. The predictors I considered largely overlapped with the variables I had used for comparing groups based on study completion (dropout)

status. For this analysis, however, I did not include a global measure of pre-stroke functional independence, as reflected by the mRS score, assuming that this construct would be jointly captured by the many variables relevant to medical history. Moreover, I did not investigate associations with measures of physical activity and social support due to the very high proportion of missing values, exceeding what is considered appropriate for imputation (267). Instead, I conducted a separate analysis with these variables in a subsample of participants, described in detail in the next section.

There were also differences in how I coded particular variables of interest, as alongside retaining maximum information, I aimed to accommodate modelling challenges, related to value distributions diverging far from normal (e.g. zero-inflated) and the presence of outliers. Specifically, for this analysis, I categorised education into four groups, in accordance with the UK schooling system: category 1 – under 11 years, which is below the current compulsory minimum for full-time education; category 2 – 11 years, reflecting compulsory duration at a General Certificates of Secondary Education (GCSE) level; category 3 – 12 and 13 years, reflecting education at Advanced Levels (A-Levels); and category 4 – above 13 years, reflecting progressing into higher education.

Further, for stroke severity as measured by the NIHSS score, I applied the categorisation presented in Table 7-3, distinguishing: no stroke signs (score of 0), minor stroke (score of 1 to 4), moderate (score of 5 to 15), and severe (score of 16 to 42) (361). Conversely, instead of implementing the clinically recognised cut-off values for BMI, I included this factor in the model as a centred, continuous variable. This was to allow more accurate modelling of a potentially non-linear relationship, with evidence suggesting that, in older age, poorer cognitive function may be associated with both low and very high BMI values (465-467). As follows, alongside a linear term, I included a quadratic term for BMI in the analysis (BMI^2).

After selecting and coding predictors for inclusion in the model, I conducted a missing value analysis, using IBM SPSS Statistics 27. I found that missing values constituted 0.8% of all values in the dataset, and related to six variables (education, BMI, history of substance abuse, smoking status, and stroke severity). A graphical inspection of the pattern of missingness indicated that the

data were likely missing completely at random. It is generally considered that in cases such as this – completely random missingness, affecting below 5% of datapoints – missing values can be ignored (267), and a complete-case analysis approach applied. However, in this study, this would have led to the exclusion of 40 participants, and thus a substantial loss of statistical power. I therefore decided to employ a multiple imputation procedure, using Bayesian analysis in Mplus (468, 469).

Multiple imputation is recommended as a method that accounts for uncertainty about the right value to impute, producing unbiased parameter estimates in a variety of missing data situations (470). A missing value is not substituted with a single value, but instead is replaced by a set of plausible values, representing a distribution of possibilities (471). Consequently, multiple datasets are generated for use in subsequent analyses, the results of which are combined for inference. For all analyses predicting class membership, I generated ten imputed datasets.

7.2.7 Additional analyses

7.2.7.1 Physical activity and social support as predictors of class membership

I included these two predictors in a multivariable analysis involving a subsample of study participants, who completed the relevant questionnaires at baseline. Given the relatively small sample size, I decided to limit the covariates to factors that seemed most essential to account for, including age, sex, education, and stroke severity. I intended to also add any variables found significant in the main logistic regression. Further, assuming that many health-related factors would likely be omitted, for this analysis I decided to account for pre-stroke mRS (as an ordinal variable with five categories). Given far from normal distributions, I categorised scores from measures of physical activity and social support based on tertiles.

7.2.7.2 Mediation analyses

An additional aim of this study was to verify whether the effects of cardiovascular risk factors described in Chapter 5 would be reproduced after redefining the outcome from “acute post-stroke cognitive function” to “pattern of longer-term cognitive change”. For this purpose, I intended to replicate the

previously developed moderated-mediation model as accurately as possible to predict class membership.

Overall, I performed all relevant statistical procedures as outlined for the previous study, with the exception of using multiple imputed datasets. As a consequence of the latter, computing bias-corrected bootstrap confidence intervals was not possible. Another important limitation was that, for APPLE, data regarding previous stroke and previous TIA had been recorded as a single variable, and therefore I could not test the individual associations of these factors with the outcome.

In summary, the structure of the investigated model would be as follows: i) latent trajectory class membership regressed on two mediators – stroke severity and prior cognitive impairment, and seven predictors – age, sex, diabetes, hypertension, vascular disease, atrial fibrillation, and previous stroke/TIA; ii) stroke severity regressed on the seven predictors and two interaction terms – between vascular disease and hypertension, and vascular disease and diabetes; iii) prior cognitive impairment regressed on the seven predictors.

Following this replication analysis, I also sought to expand the developed model. Of particular interest was inclusion of education as a predictor, the effect of which I could not account for in the previous study. Moreover, I planned to introduce any additional variables that were significantly associated with latent trajectory class membership in the main logistic regression analysis. Finally, with the model being redefined, I intended to remove potentially nonsignificant interaction terms for the prediction of stroke severity, to achieve a more parsimonious solution (368).

7.3 Summary

In this chapter, I presented the methods I had applied to explore heterogeneity in the natural history of cognitive change following stroke in the APPLE dataset. The core component of this study is the LCGA that allows to identify distinct trajectory classes within a studied population. In preparation for this stage of analysis, I: i) conducted a factor analysis to derive a latent cognitive variable based on participants' AMT-plus scores, verifying that measurement invariance had been achieved; and ii) implemented two alternative approaches for handling

missing outcome data due to dropout – based on FIML alone, and coupling FIML with selective substitution of missing values.

The end of this chapter on study methods does not mark the conclusion of a decision chain regarding approaches to analysing the APPLE data. Similarly as for the procedures I described above, selecting an optimal latent class model involves an iterative process. Its result in turn informs the specific choice of strategies in a subsequent part of the study, investigating predictors of latent class membership. Findings from these two linked investigation components are the focus of my next chapter.

Chapter 8 Trajectories of post-stroke cognitive change following stroke: A pilot study using the APPLE dataset. Part II: Results and discussion.

In the first section of this chapter, I present the results of the main and additional analyses I conducted in the APPLE dataset. As I described in Part I, this included specifying an optimal growth model based on LCGA, using a latent cognitive factor as the outcome variable, followed by identifying predictors of distinguished trajectory classes. In the second section, I discuss the clinical and research implications of recognising heterogeneity within a stroke population, with special consideration to how cognitive outcomes are understood and captured. Further, based on my experience of applying LCGA in the context of post-stroke cognition, I make recommendations for future research.

8.1 Results

8.1.1 Model selection and description of identified trajectories

8.1.1.1 Default FIML approach for handling missing outcome data

Individual trajectories of cognitive change over time are presented in Appendix 11. I initially developed linear growth models with one to six latent classes, using the default FIML approach. Table 8-1 presents characteristics relevant to optimal model selection. As a next step, I assessed the alternative quadratic growth models. The one and two-class models had poorer fit indices compared to equivalent linear growth models (BIC of 3558.7 and 3287.1, respectively), with the two-class model also presenting lower entropy (0.75). For models with three or more classes, meaningful estimates could not be obtained for all parameters (model nonidentification). Therefore, I narrowed my selection process to three linear growth models with the most favourable characteristics – with three, four and six classes. I excluded the five-class model as model fit indices were only modestly better than for the four-class alternative and, overall, out of all models classification accuracy was poorest.

At this stage, I predominantly focused on the types of distinguished trajectories, discussing the possible clinical interpretation of alternative model solutions with physicians specialising in stroke. Compared to the three-class model (Figure 8-1), the four-class model (Figure 8-2) allowed to distinguish a unique trajectory shape – latent Class 4. Unlike for the three remaining classes, in this case change over time appeared relatively constant, without a steeper period of improvement between baseline and the 1-month follow-up. In the six-class model (Figure 8-3), however, no additional, unique trajectory type was identified, with little difference in trajectory shape between Classes 1 and 2, and Classes 4 and 5. Moreover, there were relatively few participants representing Classes 1 and 6 (under 20 cases). On this basis, I selected the four-class model as the most optimal solution. Following replication of the best loglikelihood values, I concluded that local maxima had been successfully avoided.

Table 8-1 Linear growth model comparison for models developed using default FIML approach.

Model	AIC	BIC	SSA-BIC	Entropy	Proportion of sample per class	Classification accuracy
1 class	1980.7	2003.7	1984.7	———	———————	———————
2 classes	1590.4	1624.9	1596.4	0.80	59.2%, 40.8%	93.9%, 95.0%
3 classes	1335.8	1381.8	1343.8	0.85	23.3%, 53.1%, 23.6%,	94.7%, 91.5%, 92.2%
4 classes	1202.7	1260.2	1212.7	0.87	24.2%, 43.4%, 25.1%, 7.3%	93.6%, 93.2%, 90.4%, 91.5%
5 classes	1138.4	1207.5	1150.4	0.84	12.8%, 20.4%, 22.4%, 37.0%, 7.3%	85.1%, 86.4%, 93.1%, 87.6%, 92.8%
6 classes	1091.4	1172.0	1105.4	0.86	5.5%, 21.3%, 32.7%, 20.4%, 15.2%, 5.0%	96.3%, 91.3%, 89.0%, 86.5%, 90.2%, 97.1%

AIC indicates Akaike information criterion (AIC); BIC, Bayesian information criterion; SSA-BIC, sample size-adjusted Bayesian information criterion.

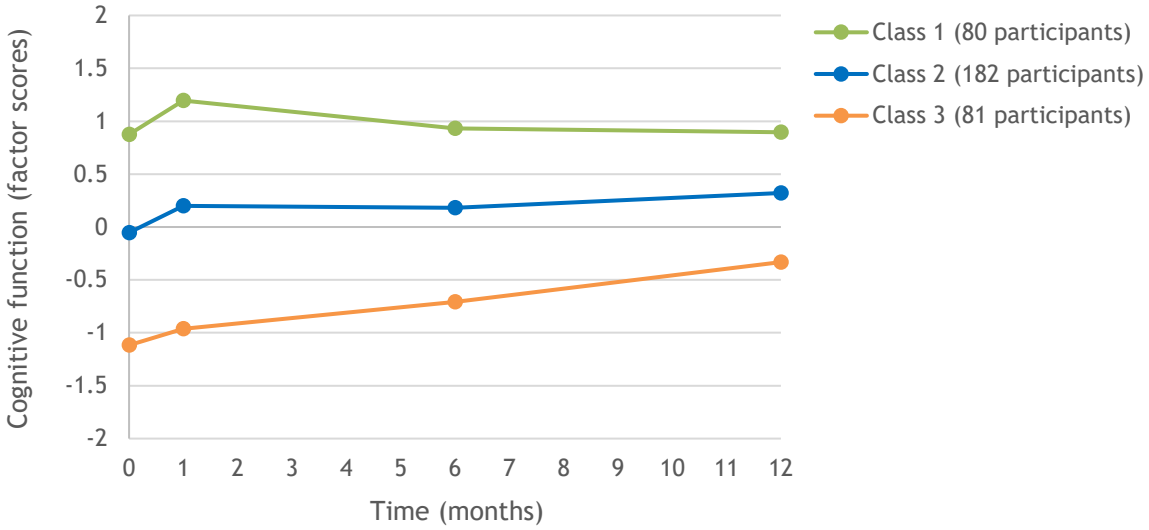


Figure 8-1 Trajectories of post-stroke cognitive change for a three-class model based on observed means, estimated using a default FIML approach.

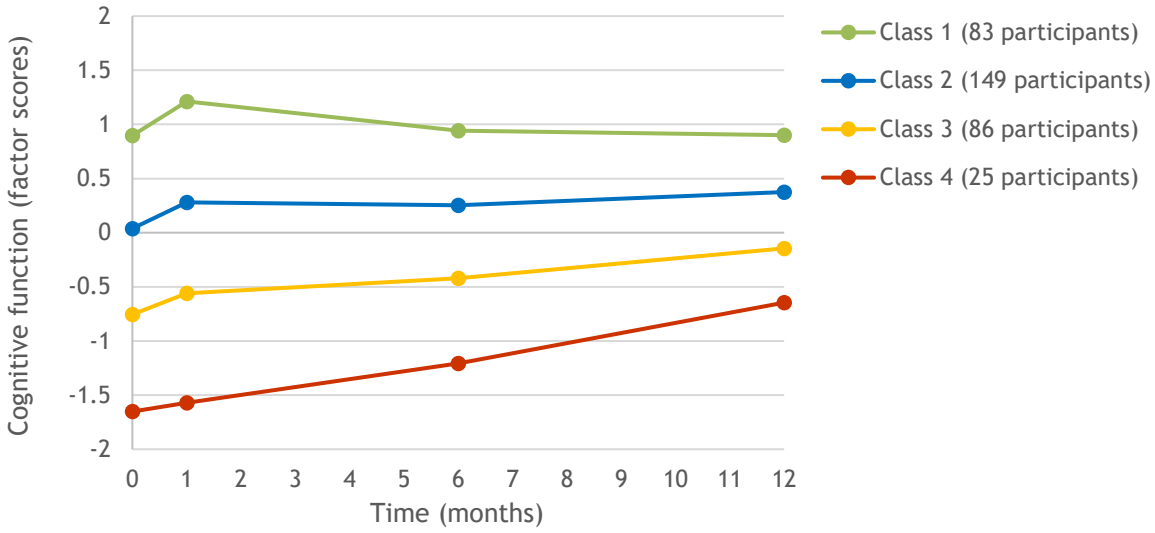


Figure 8-2 Trajectories of post-stroke cognitive change for a four-class model based on observed means, estimated using a default FIML approach.

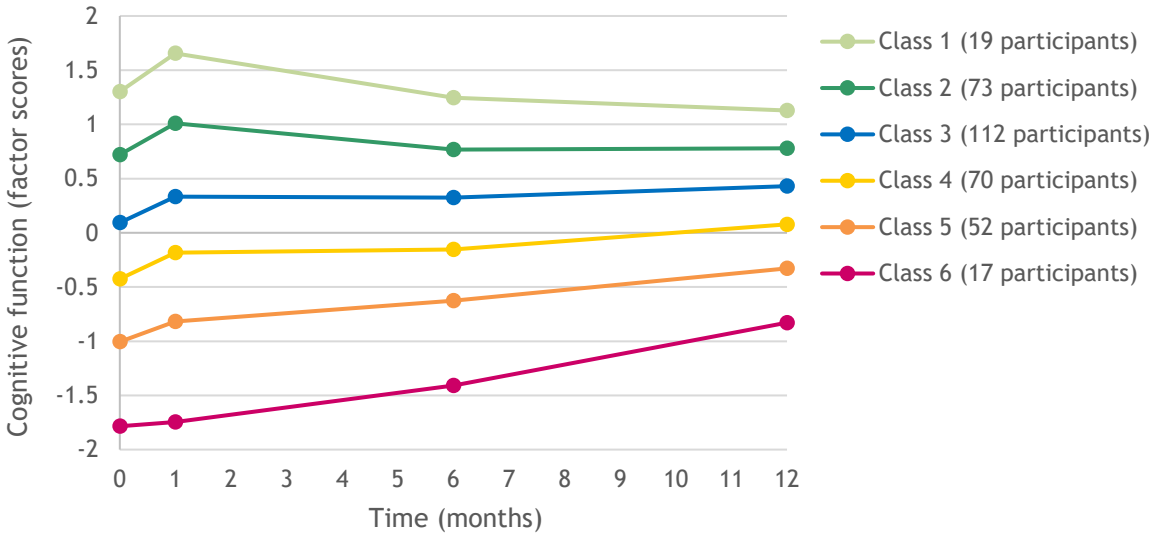


Figure 8-3 Trajectories of post-stroke cognitive change for a six-class model based on observed means, estimated using a default FIML approach.

8.1.1.2 FIML with selective substitution for handling missing outcome data

Given little support for a quadratic solution, I only developed linear growth models using the alternative approach to handling missing outcome data, combining FIML with selective substitution of missing values due to study dropout. Based on model characteristics presented in Table 8-2, the three, four, and five-class options seemed most favourable, where here it was the 6-class model for which classification accuracy was poorest.

Graphical examination indicated that the types of identified trajectories were very similar for both approaches to handling missing data. As previously, compared to the three-class model (Figure 8-4), the four-class model (Figure 8-5) led to the detection of a unique trajectory shape. For the five-class model (Figure 8-6), on the other hand, I found that despite different intercepts, the key features of trajectory shape were difficult to distinguish between Classes 1 and 2. Therefore, once again, the four-class model appeared to represent an optimal solution, which – as I verified – was not based on local maxima.

Table 8-2 Linear growth model comparison for models developed using a combined approach to handling missing outcome data.

Model	AIC	BIC	SSA-BIC	Entropy	Proportion of sample per class	Classification accuracy
1 class	2343.6	2366.6	2347.6			
2 classes	1871.0	1905.6	1877.0	0.78	41.7%, 58.3%	93.0%, 93.3%
3 classes	1573.5	1619.5	1581.5	0.85	53.1%, 18.9%, 28.0%	91.2%, 92.7%, 94.9%
4 classes	1443.4	1500.9	1453.4	0.88	25.9%, 42.9%, 8.2%, 23.0%	91.5%, 92.4%, 94.9%, 95.5%
5 classes	1372.6	1441.7	1384.6	0.88	22.7%, 7.9%, 23.9%, 39.7%, 5.8%	92.5%, 93.7%, 90.7%, 92.4%, 95.3%
6 classes	1334.4	1415.0	1348.3	0.85	15.5%, 5.8%, 20.4%, 34.4%, 5.8%, 18.1%	85.0%, 94.3%, 92.0%, 89.0%, 94.2%, 81.4%

AIC indicates Akaike information criterion (AIC); BIC, Bayesian information criterion; SSA-BIC, sample size-adjusted Bayesian information criterion.

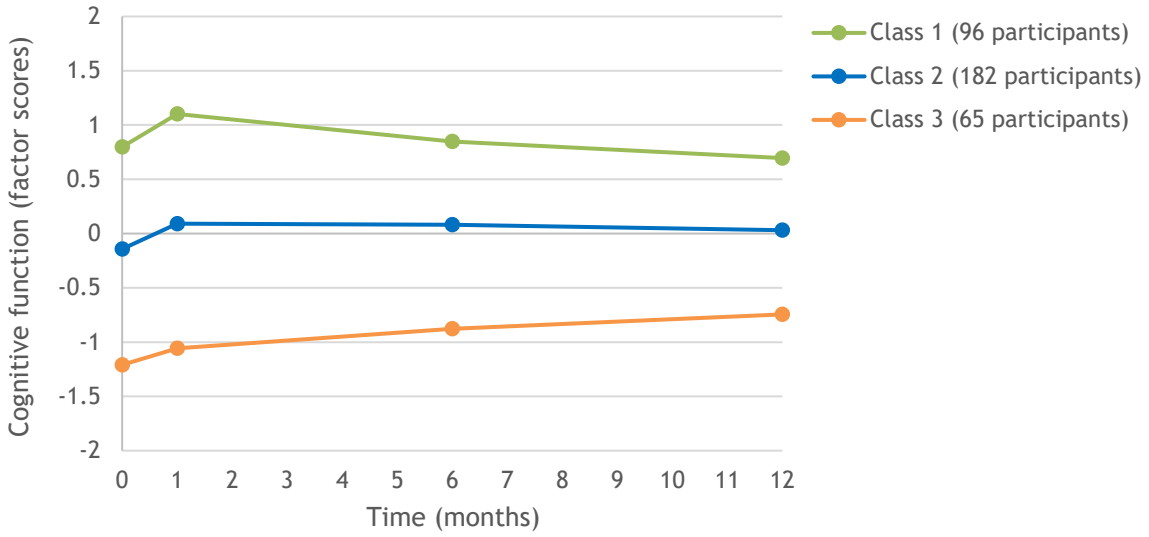


Figure 8-4 Trajectories of post-stroke cognitive change for a three-class model based on observed means, estimated using a combined approach.

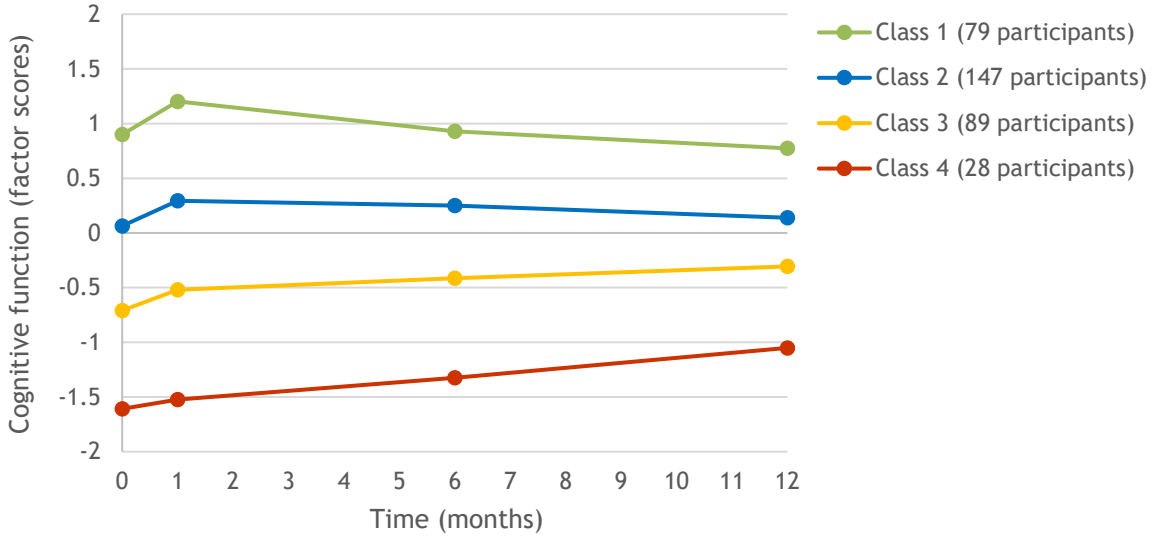


Figure 8-5 Trajectories of post-stroke cognitive change for a four-class model based on observed means, estimated using a combined approach.

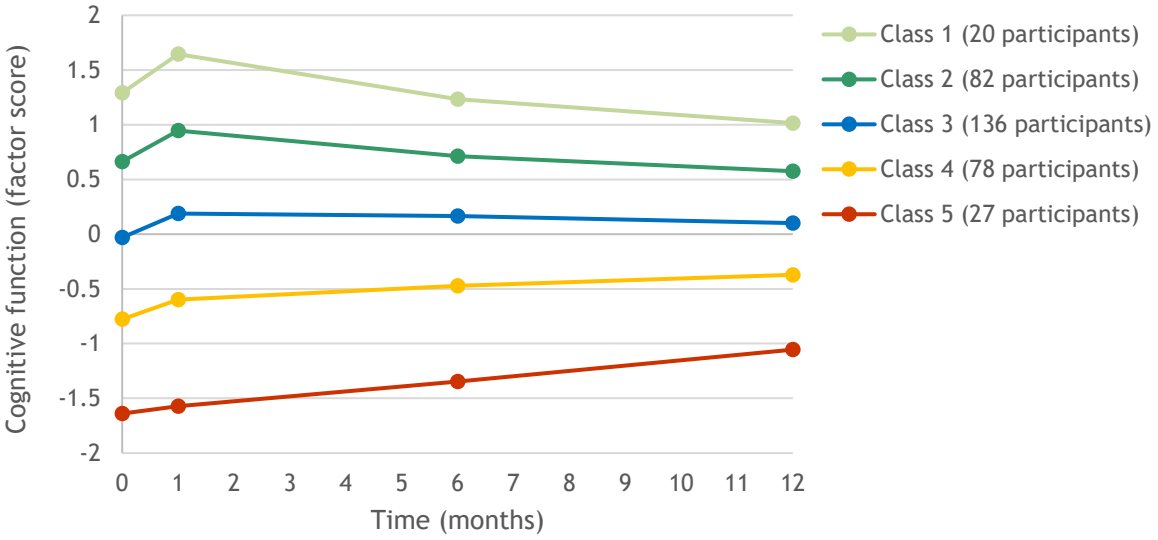


Figure 8-6 Trajectories of post-stroke cognitive change for a five-class model based on observed means, estimated using a combined approach.

8.1.1.3 Comparison of results from alternative approaches to handling missing outcome data

Table 8-3 presents the estimated intercept and slope for each class for both approaches to handling missing outcome data. Depending on the sign, a significant slope indicates either an overall improvement or decline in cognitive function. However, it is important to note that this estimate represents an average across all timepoints, and as such, cannot fully reflect a trajectory shape where change is not relatively constant. This was particularly relevant in case of Class 1, where the direction of slope changed. In view of this, in my description of identified trajectories, I considered both estimated parameters and graphical representations, assuming that real values were likely to lie at an intermediate point between results obtained using the two different approaches to handling missing outcome data.

Table 8-3 Estimated class characteristics for four-class models by approach to handling missing outcome data.

	Default FIML approach	Combined approach
Class 1		
Intercept	1.021	1.016
Slope	-0.002*	-0.004
Class 2		
Intercept	0.097	0.151
Slope	0.006*	<0.001
Class 3		
Intercept	-0.730	-0.665
Slope	0.011*	0.007*
Class 4		
Intercept	-1.657	-1.613
Slope	0.017*	0.010*

*significant at $p < 0.05$

FIML indicates full information maximum likelihood.

As follows, I concluded that Class 1 was characterised by high cognitive function soon after stroke, which improved over following weeks, and thereafter declined. Class 2 presented with some cognitive problems soon after stroke, followed by a period of improvement, after which cognitive function remained relatively stable. Class 3 was characterised by comparatively poor initial cognitive function, which after a stage of steeper improvement, continued to improve at a slower rate. Class 4 experienced severe cognitive impairment soon

after stroke, which was followed by improvement in cognitive function at a relatively constant rate.

Although in my interpretation of the selected four-class solution I considered results from both models, only one model could be chosen for assigning final class membership. For this purpose, I selected the model I developed by combining FIML with selective substitution of missing outcome values. This decision was guided by two considerations. Firstly, using the default approach, from the 6-month timepoint trajectories of all classes showed a positive change – slower decline for Class 1 and greater improvement for the three remaining classes (see Figure 8-2). From a clinical perspective, it seemed implausible that this reflected a real change in cognitive function and, as follows, it could be suspected that this effect resulted from increased sample bias after participants with poorer cognition (relative to their assigned class) dropped out of the study. Secondly, for the default approach, classification accuracy was slightly poorer.

8.1.2 Prediction of class membership

Descriptive statistics for predictors of interest by assigned latent class membership are provided in Table 8-4. I presented the variables to reflect how they were entered in the analyses (as binary, categorical or continuous), with the exception of BMI, which while remaining continuous, I centred for inclusion in developed models. To facilitate conveying of results, I designated the trajectory classes based on the represented overall, relative level of cognitive function and unique trajectory features: Class 1 – high – declining, Class 2 – mid-high – stable, Class 3 – mid-low – slowing improvement, Class 4 – low – constant improvement.

Table 8-4 Descriptive statistics for predictors of interest by class membership.

	High - declining (N = 79)	Mid-high - stable (N = 147)	Mid-low - slowing improvement (N = 89)	Low - constant improvement (N = 28)
Age, Mean (SD)	66.2 (11.5)	68.2 (13.5)	71.6 (12.9)	75.6 (9.3)
Missing	0	0	0	0
Sex				
Female, N (%)	30/79 (38.0%)	66/147 (44.9%)	43/89 (48.3%)	14/28 (50.0%)
Education				
<11 years	10/74 (13.5%)	46/136 (33.8%)	40/82 (48.8%)	11/22 (50.0%)
11 years	11/74 (14.9%)	39/136 (28.7%)	30/82 (36.6%)	8/22 (36.4%)
12-13 years	23/74 (29.1%)	30/136 (22.1%)	6/82 (7.3%)	3/22 (13.6%)
>13 years	30/74 (40.5%)	21/136 (15.4%)	6/82 (7.3%)	0/22 (0.0%)
BMI, Mean (SD)	29.1 (6.4)	28.3 (5.4)	26.5 (6.1)	26.6 (5.2)
Missing	1	4	2	1
Smoking status				
Never, N (%)	37/79 (46.8%)	49/146 (33.6%)	31/89 (34.8%)	10/27 (37.0%)
Former, N (%)	28/79 (35.4%)	65/146 (44.5%)	37/89 (41.6%)	10/27)37.0%)
Current, N (%)	14/79 (17.7%)	32/146 (21.9%)	21/89 (23.6%)	7/27 (26.0%)
History of substance abuse, N (%)	8/79 (10.1%)	20/145 (13.8%)	7/89 (7.9%)	2/28 (7.1%)
Diabetes, N (%)	20/79 (25.3%)	35/147 (23.8%)	21/89 (23.6%)	9/28 (32.1%)
Hypertension, N (%)	41/79 (51.9%)	78/147 (53.1%)	48/89 (53.9%)	16/27 (59.3%)
Vascular disease, N (%)	19/79 (24.1%)	45/147 (30.6%)	21/89 (23.6%)	13/28 (46.4%)
Heart failure, N (%)	5/79 (6.3%)	10/137 (6.8%)	7/89 (7.9%)	4/28 (14.3%)
Atrial fibrillation, N (%)	9/79 (11.4%)	24/147 (16.3%)	16/89 (18.0%)	7/28 (25.0%)
Previous stroke/TIA, N (%)	13/79 (16.5%)	37/147 (25.2%)	23/89 (25.8%)	12/28 (42.9%)
Renal disease, N (%)	7/79 (8.9%)	16/147 (10.9%)	12/89 (13.5%)	6/28 (21.4%)
Prior cognitive impairment, N (%)	2/79 (2.5%)	4/147 (2.7%)	9/89 (10.1%)	11/28 (39.3%)
History of mood disorder, N (%)	19/79 (24.1%)	39/147 (26.5%)	25/89 (28.1%)	8/28 (28.6%)
Physical activity				
1st tertile, N (%)	12/38 (31.6%)	28/79 (35.4%)	16/39 (41.0%)	5/9 (55.6%)
2nd tertile, N (%)	14/38 (36.8%)	28/79 (35.4%)	13/39 (33.3%)	1/9 (11.1%)
3rd tertile, N (%)	12/38 (31.6%)	23/79 (29.2%)	10/39 (25.7%)	3/9 (33.3%)
Social support				
1st tertile, N (%)	10/38 (26.3%)	27/79 (34.2%)	13/39 (33.3%)	6/9 (66.7%)
2nd tertile, N (%)	9/38 (23.7%)	21/79 (26.6%)	7/39 (18.0%)	0/9 (0.0%)
3rd tertile, N (%)	19/38 (50.0%)	31/79 (39.2%)	19/39 (48.7%)	3/9 (33.3%)
Stroke severity (NIHSS)				
No stroke signs, N (%)	23/79 (29.1%)	40/145 (27.6%)	19/89 (21.3%)	0/28 (0.0%)
Mild, N (%)	47/79 (59.5%)	76/145 (52.4%)	46/89 (51.7%)	10/28 (35.7%)
Moderate, N (%)	9/79 (11.4%)	29/145 (20.0%)	21/89 (23.6%)	11/28 (39.3%)
Severe, N (%)	0/79 (0.0%)	0/145 (0.0%)	3/89 (3.4%)	7/28 (25.0%)

BMI indicates body mass index; SD, standard deviation; TIA, transient ischaemic attack.

8.1.2.1 Results of multinomial logistic regression analysis

The results of this analysis, central to the prognostic objective of this study, are presented in Table 8-5. I chose the mid-high – stable class as an initial point of reference for all remaining latent trajectory classes, as one that: included most participants, had an intermediate intercept estimate, showed relatively limited change in cognitive function over time, and appeared closest to the post-stroke recovery pattern traditionally described in existing literature – initial (spontaneous) improvement, followed by a plateau period (472-474).

Compared to this class, I found that representing the high – declining trajectory was nearly two times more likely by moving up one category of education, which constituted the only significant finding for this comparison. Conversely, with moving up one education category, the likelihood of representing the mid-low – slowing improvement class decreased by over a third. Compared to the mid-high stable class, participants here were also 3.5 times less likely to have a history of substance abuse, while being nearly 5 times more likely to have a history of cognitive impairment.

Further, associations with BMI suggested a possible non-linear relationship, where the likelihood of being in the mid-low – slowing improvement class decreased with rising BMI until approximately a value of 34.5 (obesity), after which it began to increase. In other words, participants in this class were more likely to have both relatively low and very high BMI, while high-mid-range values were associated with the mid-high stable class. I moreover observed a trend ($p = 0.056$), suggesting that as stroke increased in severity from one category to the next, participants were around 50% more likely to represent the mid-low – slowing improvement trajectory.

For the third comparison with the mid-high – stable class as a reference, the likelihood of representing the low – constant improvement trajectory increased nearly by one-tenth with every one-year increase in age, increased 27 times with a history of prior cognitive impairment, and 12 times with moving up a category of stroke severity. I also noted a trend ($p = 0.055$), indicating that participants were around two times less likely to belong to this class with the overall poorest level of cognitive function as education increased by one category.

Table 8-5 Results of multinomial logistic regression analysis identifying predictors of latent trajectory class membership.

	Mid-high - stable[a] vs. high - declining	Mid-high - stable[a] vs. mid-low – slowing improvement	Mid-high – stable[a] vs. low – constant improvement	High – declining[a] vs. mid-low – slowing improvement	High – declining[a] vs. low – constant improvement	Mid-low – slowing improvement[a] vs. low – constant improvement
Age	0.99 (0.96, 1.02)	1.01 (0.98, 1.04)	1.08 (1.01, 1.15)*	1.02 (0.99, 1.05)	1.09 (1.02, 1.16)**	1.06 (1.00, 1.13)*
Sex	1.02 (0.53, 1.98)	1.32 (0.71, 2.42)	1.53 (0.51, 4.61)	1.29 (0.59, 2.81)	1.50 (0.45, 5.03)	1.16 (0.40, 3.43)
Education	1.90 (1.38, 2.61)**	0.64 (0.46, 0.88)**	0.47 (0.22, 1.02)	0.34 (0.23, 0.50)**	0.25 (0.11, 0.56)**	0.74 (0.34, 1.60)
BMI	0.99 (0.93, 1.06)	0.92 (0.87, 0.98)*	0.98 (0.86, 1.11)	0.93 (0.87, 1.00)*	0.99 (0.87, 1.12)	1.06 (0.93, 1.20)
BMI^2	1.00 (1.00, 1.01)	1.01 (1.00, 1.01)*	0.99 (0.98, 1.01)	1.00 (1.00, 1.01)	0.99 (0.98, 1.01)	0.99 (0.97, 1.00)
Smoking status	0.75 (0.45, 1.23)	1.15 (0.76, 1.74)	1.26 (0.61, 2.59)	1.54 (0.89, 2.66)	1.69 (0.74, 3.83)	1.10 (0.54, 2.23)
History of substance abuse	1.05 (0.34, 3.22)	0.28 (0.10, 0.79)*	0.30 (0.04, 2.54)	0.27 (0.07, 0.96)*	0.28 (0.03, 2.80)	1.06 (0.13, 8.75)
Diabetes	1.31 (0.64, 2.69)	1.02 (0.49, 2.10)	1.32 (0.37, 4.68)	0.78 (0.34, 1.77)	1.00 (0.26, 3.86)	1.29 (0.37, 4.51)
Hypertension	0.98 (0.53, 1.80)	1.10 (0.60, 2.02)	1.13 (0.41, 3.14)	1.12 (0.55, 2.31)	1.16 (0.39, 3.47)	1.03 (0.38, 2.80)
Vascular disease	0.92 (0.45, 1.88)	0.62 (0.29, 1.32)	1.00 (0.30, 3.41)	0.68 (0.27, 1.68)	1.09 (0.29, 4.16)	1.62 (0.49, 5.36)
Heart failure	1.51 (0.41, 5.53)	1.27 (0.37, 4.34)	1.83 (0.34, 9.87)	0.84 (0.18, 3.97)	1.22 (0.19, 7.83)	1.44 (0.27, 7.67)
Atrial fibrillation	1.00 (0.39, 2.61)	0.90 (0.42, 1.93)	0.43 (0.10, 1.90)	0.90 (0.31, 2.56)	0.42 (0.08, 2.23)	0.47 (0.11, 2.04)
Previous stroke/TIA	0.58 (0.26, 1.29)	0.95 (0.46, 1.96)	1.83 (0.64, 5.23)	1.63 (0.63, 4.22)	3.16 (0.95, 10.50)	1.93 (0.69. 5.41)
Renal disease	0.92 (0.28, 2.98)	1.13 (0.47, 2.74)	1.05 (0.32, 3.45)	1.24 (0.32, 4.82)	1.14 (0.23, 5.62)	0.93 (0.28, 3.05)
Prior cognitive impairment	1.07 (0.20, 5.89)	4.85 (1.24, 18.92)*	26.8 (5.39, 133.02)**	4.51 (0.79, 25.85)	24.94 (3.53, 176.05)**	5.53 (1.35, 22.71)*
History of mood disorder	1.01 (0.46, 2.24)	1.21 (0.62, 2.38)	0.53 (0.10, 2.67)	1.20 (0.51, 2.84)	0.52 (0.09, 2.98)	0.43 (0.09, 2.10)
Stroke severity	0.75 (0.48, 1.16)	1.51 (0.99, 2.31)	11.94 (4.64, 30.7)**	2.02 (1.21, 3.38)**	15.96 (5.99, 42.5)**	7.90 (3.07, 20.35)**

[a]reference class; *significant at $p < 0.05$; **significant at $p < 0.01$

BMI indicates body mass index; TIA, transient ischaemic attack.

With the high – declining class as a reference, findings were in most part similar. Representing the mid-low – slowing improvement class was approximately three times less likely for both education higher by one category and a history of substance abuse. I also observed an association where for every 1 unit increase in BMI, the likelihood for cognitive change to follow this trajectory decreased by 7%. At the same time, belonging to the mid-low – slowing improvement class was two times more likely for every one-category increase in stroke severity.

On comparing the high -declining to the low – constant improvement class, I observed that representing the latter trajectory was more likely by nearly one-tenth with every one-year increase in age, 25 times more likely with a history of cognitive impairment, and 16 times with moving up a category of stroke severity. Conversely, moving up one category of education was associated with being four times less likely to belong to the low – constant improvement class. I moreover observed a trend ($p = 0.061$), suggesting the participants in this class were approximately three times more likely to have had a previous stroke or TIA.

In the final comparison, with the mid-low – slowing improvement class serving as a reference, I found that the likelihood of representing the low – constant improvement trajectory increased 0.06 times with every one-year increase in age, 5.5 times with a history of cognitive impairment, and nearly 8 times with a one-category increase in stroke severity.

8.1.3 Results of additional analyses

8.1.3.1 Associations of latent class membership with physical activity and social support

This analysis involved a subsample of 165 participants with data on self-reported physical activity and social support, including: 38 participants in the high – declining class, 79 in the mid-high – stable class, 39 in the mid-low – slowing improvement class, and 9 in the low – constant improvement class. Considering results of the main analysis alongside my initial assumptions, I accounted for the following predictors: age, sex, education, BMI (linear and quadratic term), pre-stroke mRS, history of substance abuse, previous stroke/TIA, prior cognitive impairment, and stroke severity.

However, even with a limited number of variables, I recognised that low statistical power would present a major issue for this analysis. To at least partially ameliorate this limitation, I decided to combine the two higher-function classes and the two lower-function classes. In addition to relative similarity in the overall level of cognitive function, this seemed justified by the observed pattern of change, with the lower-function classes presenting continuing improvement, unlike the other two class. Consequently, for this part of the study I conducted a binary logistic regression.

I found no association between measures of physical activity and social support and assignment to a lower-function class (OR = 1.02, 95% CI: 0.73 to 1.42; and OR = 1.02, 95% CI: 0.76 to 1.36, respectively). Among significant predictors were only education (OR = 0.66, 95% CI: 0.52 to 0.85) and stroke severity (OR = 1.48, 95% CI:1.05 to 2.09).

8.1.3.2 Mediation analyses

Similarly as described above, in view of low statistical power, I used a binary outcome, differentiating between the two higher-function latent classes and two lower-function classes. The first analysis was aimed at replicating the moderated mediation model described in Chapter 5. Here, however, obtained estimates suggested poor model fit. Following modification recommendations, which can be requested as part of the software output, I therefore included an additional path, with stroke severity regressed on prior cognitive impairment. The resulting mean fit index estimates suggested very good model fit: CFI = 1.00, TLI = 1.23, RMSEA < 0.01, SRMR = 0.03.

For class membership, I found a direct association only with the two mediators: stroke severity (coefficient = 0.218, 95% CI: 0.001 to 0.435) and prior cognitive impairment (coefficient = 0.580, 95% CI: 0.241 to 0.918). In relation to mediator predictors, stroke severity was inversely associated with age (coefficient = −0.019, 95% CI: −0.036 to −0.002), and positively associated with atrial fibrillation (coefficient = 0.515, 95% CI: 0.129 to 0.902) and prior cognitive impairment (coefficient = 0.354, 95% CI: 0.143 to 0.566). Prior cognitive impairment was significantly predicted only by age (coefficient = 0.036, 95% CI: 0.006 to 0.066). Although, there also appeared to be a trend for an association with diabetes (coefficient = 0.601, 95% CI: −0.091 to 1.293, p = 0.089).

In view of the observed direct associations, as a next step, I tested plausible indirect effects. Of these, only two were significant. Representing a lower-function class was associated with prior cognitive impairment through greater stroke severity (coefficient = 0.077, 95% CI: 0.013 to 0.142); and with age, through increased likelihood of prior cognitive impairment (coefficient = 0.021, 95% CI: 0.001 to 0.041). I further found a trend for two additional indirect effects, where age decreased the likelihood of belonging to a lower-function class through reduced stroke severity (coefficient = −0.004, 95% CI: −0.009 to 0.001, p = 0.086), while atrial fibrillation increased this likelihood through greater stroke severity (coefficient = 0.112, 95% CI: −0.016 to 0.241, p = 0.087).

In the second analysis, I additionally included education, BMI, and history of substance abuse among predictors. To develop a final model, I removed insignificant interaction terms one by one in order of descending p-value, which here also related to a quadratic term for BMI. However, removing the latter as a predictor of class membership resulted in a deterioration of model fit, and therefore I retained this term. Fit indices for the resulting final model, presented in Figure 8-7, indicated very good model fit: CFI = 1.00, TLI = 1.20, RMSEA < 0.01, SRMR = 0.02.

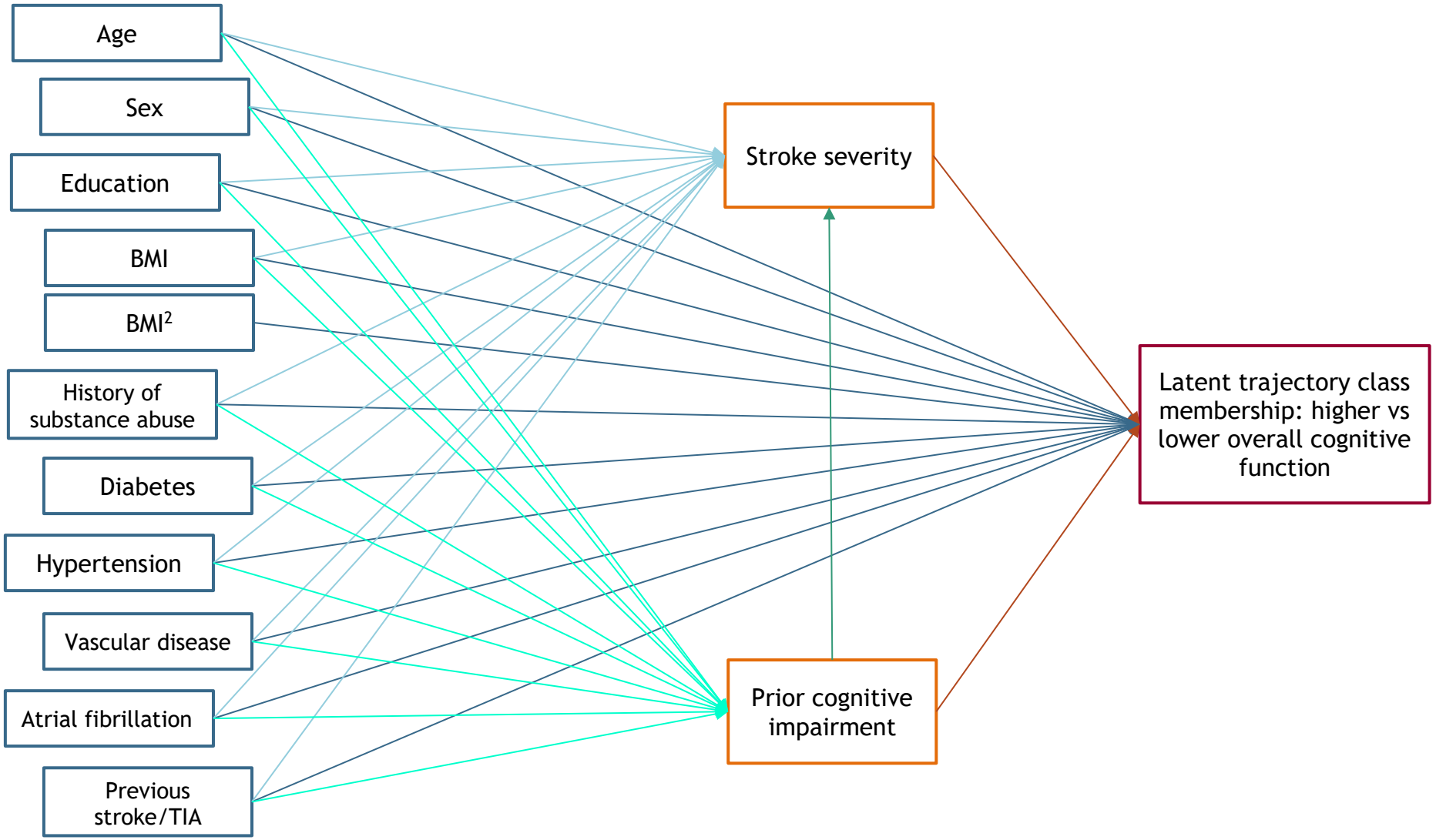


Figure 8-7 Conceptual diagram of the final mediation model with two parallel mediators for prediction of latent trajectory class membership.

Direct associations between predictors, mediators and class membership are presented in Table 8-6. As in the previous analysis, representing a lower-function class was more likely with greater stroke severity and prior cognitive impairment. This likelihood decreased, however, with education and history of substance abuse. I further observed a trend for an inverse association between BMI and lower cognitive function (p = 0.054). Although, if the estimate for the quadratic term of BMI indeed represented a true effect (p = 0.113), this would indicate a change in the direction of the association for very high BMI values – the likelihood of belonging to a lower-function class would start to increase from a BMI of approximately 36.5.

The effects of predictors on mediators mostly replicated findings from the previous analysis. Reduced stroke severity was associated with age, while greater severity – with atrial fibrillation and prior cognitive impairment. The latter was significantly associated with age only, although I also observed a trend for an association with history of substance abuse (p = 0.082).

Table 8-6 Direct associations between predictors of interest and stroke severity, prior cognitive impairment, and class membership; comparing combined two higher-function classes (reference) with two lower-function classes.

	Unstandardised coefficients (95% CI)		
	Stroke severity	**Prior cognitive impairment**	**Class membership**
Stroke severity	————	————	0.264 (0.047, 0.482)*
Prior cognitive impairment	0.330 (0.120, 0.506)**	————	0.594 (0.295, 0.893)**
Age	-0.021 (-0.042, -0.001)*	0.048 (0.006, 0.090)*	-0.013 (-0.042, 0.016)
Sex	-0.125 (-0.427, 0.176)	-0.132 (-0.668, 0.404)	0.223 (-0.224, 0.669)
Education	0.068 (-0.086, 0.222)	-0.094 (-0.393, 0.205)	-0.404 (-0.642, -0.204)**
BMI	-0.023 (-0.057, 0.011)	0.009 (-0.058, 0.075)	-0.052 (-0.104, 0.001)
BMI^2	————	————	0.003 (-0.001, 0.006)
History of substance abuse	0.077 (-0.450, 0.605)	0.859 (-0.110, 1.828)	-1.162 (-2.030, -0.294)**
Diabetes	0.143 (-0.250, 0.535)	0.426 (-0.167, 1.019)	-0.200 (-0.705, 0.305)
Hypertension	0.100 (-0.219, 0.420)	-0.090 (-0.715, 0.535)	0.109 (-0.367, 0.585)
Vascular disease	-0.179 (-0.527, 170)	0.381 (-0.176, 0.938)	-0.345 (-0.824, 0.134)
Atrial fibrillation	0.571 (0.187, 0.955)**	-0.590 (-1.371, 0.192)	0.221 (-0.391, 0.833)
Previous stroke/TIA	0.187 (-0.171, 0.544)	0.140 (-0.380, 0.659)	0.080 (-0.392, 0.552)

*significant at p < 0.05; significant at p < 0.01

BMI indicates body mass index; CI, confidence interval; TIA, transient ischaemic attack.

Regarding indirect effects, I found that representing a lower-function class was significantly associated with atrial fibrillation and prior cognitive impairment through greater stroke severity (coefficient = 0.151, 95% CI: 0.005 to 0.297; and coefficient = 0.087, 95% CI: 0.023 to 0.151, respectively), as well as with age through increased risk of prior cognitive impairment. There was also some indication for possible opposing effects of age through stroke severity. As age was associated with reduced stroke severity, it in turn appeared to decrease the likelihood of belonging to a lower-function class (coefficient = −0.006, 95% CI: −0.012 to 0.001, p = 0.073). At the same time, age was associated with an increased risk of prior cognitive impairment, in turn linked to greater stroke severity, and ultimately – with lower cognitive function (coefficient = 0.004, 95% CI: 0.000 to 0.009, p = 0.079). Finally, I noted a potential indirect effect of history of substance abuse on lower-function class membership through an increased risk of prior cognitive impairment (coefficient = 0.510, 95% CI: −0.080 to 1.100, p = 0.090).

8.2 Discussion

Findings from my study confirm existing concepts. Following stroke, most individuals experience an initial period of cognitive improvement, after which function may either continue to improve, remain relatively stable or decline. However, with improvement associated with overall more severe impairment, and decline with high cognitive function, an interpretation of what constitutes ‘good’ post-stroke cognitive outcome remains uncertain.

Based on the APPLE dataset, I identified four distinct trajectory classes for post-stroke cognitive change over a one-year period. Key distinguishing features of trajectory shape related to the intercepts and direction of change following the 1-month assessment. Although regarding the latter, it is relevant to note that the uniqueness of the mid-high – stable class was rather due to the relative lack of change. The low – constant improvement class can also be set apart on account of a feature indicated in its designation – having the only trajectory representing nearly linear growth. This may be due to an initially slower recovery rate, which has been previously associated with particularly severe acute post-stroke impairment (475-477). Considering how trajectories related to one another, it is moreover important to highlight that none crossed over, that

is, the order of classes according to level of cognitive performance prevailed throughout the duration of follow-up.

Through investigating what factors characterised the identified trajectory classes, I obtained some unexpected results. Firstly, higher overall cognitive function, as represented by the first two trajectory classes, was associated with a history of substance abuse. This is in contrast to an extensive body of evidence indicating that alcohol and drug abuse increase the risk of neuropathology and cognitive deficits, with some detrimental effects likely to persist even after prolonged abstinence (478-481). Interestingly, results of the mediation analysis provided some support for such an association – I observed a trend, where history of substance abuse appeared to increase the risk of prior cognitive impairment. In view of this, it seems plausible that substance abuse was identified as a distinctive profile feature of participants with overall higher cognitive function not on the basis of a causal relationship, but rather due to covariation.

On one hand, substance abuse may have been one of the crucial predisposing factors for stroke among individuals who were comparatively young and had low comorbidity burden, and thus in some ways were less likely to present with cognitive impairment at this stage (482). On the other, substance abuse is associated with even several times higher mortality rates compared to the age-matched general population (483, 484). Recognising death as a competing risk for cognitive impairment may explain why fewer participants with a history of substance abuse represented the two lower-function trajectory classes, characterised by older age, higher prevalence of most diseases, and more severe strokes.

Another controversial finding relates to the potential duality of the indirect effect of age, increasing the likelihood of belonging to a lower-function class through increased risk of prior cognitive impairment, while decreasing this likelihood through reduced stroke severity. What was more, it seemed that age could also indirectly contribute to increased stroke severity through its effect on prior cognitive impairment. Given limitations of the analysis, and the discussed associations appearing as trends, these findings should be interpreted with

caution. Nonetheless, it seems worth considering why age could be associated with reduced stroke severity.

One interpretation of this effect, which I mentioned above in the context of alcohol abuse, relates to the presence of survival bias. Older people may be more likely to die following a severe stroke, and therefore as a group will be underrepresented in research such as this (485). As an alternative explanation, studies suggest that stroke at a younger age is associated with a greater risk of developing space occupying oedema, including malignant middle cerebral artery infarction, which is associated with severe presentation and increased fatality (486, 487). Older people are less susceptible to this condition, argued to be due to more advanced cerebral atrophy, which affords potential compensatory space within the intracranial cavity.

This phenomenon adds to other examples described in Chapter 5, indicating that in specific circumstances pathological processes can lead to more favourable outcomes. In this study, however, I did not observe a significant effect of vascular disease on reduced stroke severity (either conditional or unconditional), while an association with previous TIA alone could not be assessed. It is yet possible that, to some extent, the observed effect of age on stroke severity captured the cumulative influence of endogenous adaptive mechanisms, potentially developing with age-related progression of cardiovascular diseases.

The conclusions from Chapter 5 also allowed me to anticipate the overall lack of direct effects of cardiovascular risk factors on cognitive trajectories in the present analyses. One exception was a trend for participants with the poorest overall level of cognitive function to be more likely to have had a previous stroke or TIA as compared to the high – declining class. This is in line with existing research evidence, similarly as the observed associations between representing one of the two lower-function classes and older age, lower education, greater stroke severity, prior cognitive impairment, and both lower and very high BMI values (23, 145, 465, 467).

What is, however, puzzling, is that these factors would also typically be associated with cognitive decline, whereas the two classes characterised by them showed improvement over time. Conversely, it was the class with the

highest education and lowest prevalence of clinical risk factors (with the exception of history of alcohol abuse) that showed a decline in cognitive function after the 1-month assessment.

One explanation for this may be that trajectories of post-stroke cognitive change are influenced by two co-occurring processes, exerting opposing effects. The first involves ongoing neurodegeneration, while the second entails recovery from stroke-induced damage. For participants who had less severe strokes and were relatively unaffected by acute cognitive impairment, the scope for longer-term recovery would have been limited. This in turn could have allowed for a more pronounced effect of a neurodegenerative process, resulting from pre-existing risk factors that predisposed to stroke, as well as neuropathological sequala of the stroke itself. Given the comparatively overall high level of cognitive function of participants who represented this trajectory, it seems that this process was captured at an early stage, or at least the manifestation of it was. The latter could have been delayed through the impact of cognitive reserve (488), presumably associated with the high level of education characterising this class.

The opposite would apply to participants with overall poor cognitive function, for whom the trajectory slope appeared to predominantly reflect recovery. Nonetheless, the impact of neurodegeneration seems also apparent in this case, particularly on account of the intercept. Although poor initial cognitive function could in part be accounted for by on average greater stroke severity, findings based on medical history suggest that prior cognitive impairment was also a key contributing factor here.

In understanding the meaning of distinct trajectory shapes, it seems moreover important to emphasise that the present study offers only a “snapshot” of cognitive changes occurring over a lifespan. Trajectory class membership is plausibly to some degree fluid and dependent on the point in time at which an individual is observed. To illustrate this argument, it is possible that with age and the consequent accumulation of health-related risk factors, a person who once represented the mid-high – stable class will experience a subsequent stroke, initiating change along a low – constant improvement trajectory.

8.2.1 Clinical implications

A similar message can be emphasised as in Chapter 5 – for some stroke survivors, the risk of cognitive decline and potential future impairment may be easily underestimated. Based on my findings, this concern particularly relates to individuals representing a high - declining trajectory of cognitive change. Compared to other stroke survivors, this group's profile seems to be characterised by younger age, higher education, fewer comorbidities, a less severe index stroke, and high acute cognition. As follows, they can understandably be assumed to make a very good recovery and maintain high cognitive function. However, even where gradual decline does not amount to a diagnosis of cognitive impairment, it can be experienced by the individual as a severe loss, compromising quality of life.

Although in line with current clinical guidelines (489), all stroke survivors are to be supported in management of cardiovascular risk factors, there remains a key issue of whether anything more can be done to promote cognitive resilience and improvement. In this context, recognising distinct subpopulations among stroke survivors and understanding their characteristics may be an important step towards tailoring effective interventions to best suit individual profiles.

8.2.2 Research implications

Insights from this study encourage a revaluation of the currently dominant approach to how post-stroke cognitive outcome is defined and measured. Unlike, for example, recurrent stroke or death, cognitive function is not inherently binary, nor does it constitute a single event. Yet, in many studies in this area of research, it is treated as such. This does not imply that categorising individuals at a specific point in time following stroke into cognitively intact and cognitively impaired lacks real-world meaning or implications, or that this approach will not allow identification of relevant determinants of cognitive status. It is, however, undeniably leading to a loss of important information.

In a clinical context, the aim of prediction research is to identify individuals at risk of an unfavourable outcome so that risk can be addressed – either in the present, or in the future, for example, once an intervention can be implemented

in routine practice following successful trials (96). However, with a binary outcome assessed at one timepoint, the risk of cognitive decline cannot be recognised, and thus addressed in those whose deficits have not yet reached the threshold for diagnosis of cognitive impairment. Similarly, it will not be possible to identify, and thus attempt to influence factors that differentiate between stroke survivors who partially recover from severe cognitive impairment and those who either maintain the same poor level of function or even deteriorate.

Regardless of design, one adjustment that can be easily introduced to many studies is use of a continuous cognitive outcome. Understandably, both researchers and clinicians may consider interpretation of findings in relation to a formal diagnosis, such as mild cognitive impairment and dementia, more intuitive and useful. This, however, does not preclude from conducting an additional analysis with a continuous cognitive score for comparison of results, to verify whether any relevant effects have been missed.

Another option worth considering, although entailing a higher level of complexity, is latent class analysis, where unlike in this study, subpopulations are distinguished based on co-occurrence of certain individual characteristics (demographics, medical conditions, lifestyle factors) (490). This allows to determine distinct profiles within a heterogenous population, such as that of stroke survivors, which can then be introduced in a model to predict cognitive outcome. Yet preferably, where resources permit, future research into post-stroke cognitive outcomes should focus on longitudinal study designs with multiple assessment timepoints.

In this context, it is important to discuss one of the key challenges of longitudinal research – handling of missing outcome data due to dropout. Similarly as reported by other researchers (e.g., 168, 491, 492), on comparing participants who remained in the study to those who dropped out, I found that the latter presented more risk factors for cognitive decline. As follows, it seems that this type of missingness is likely not to be random. Such comparisons are, however, rarely conducted and most studies adopt a complete-case analysis approach. As individuals with more severe difficulties are likely to be underrepresented, this can lead to biased estimates.

Once missing data appear, there is no ideal method for handling this issue, although some techniques are considered more favourable than others (493). With this in mind, it is recommended to test more than one approach, using different assumptions about missingness, and compare the obtained results (i.e. conduct a sensitivity analysis). Observations from this comparison can then be considered for the interpretation of findings and overall study conclusions, offering a more realistic view of the investigated effects themselves, or at least the likely degree of bias that should be taken into account.

8.2.3 Strengths and limitations

The broad inclusion criteria for recruitment into APPLE, as well as the option to obtain consent through proxy, constituted key prerequisites for the participant sample to be representative of a real-world stroke case-mix. I aimed to preserve this advantage in the present study, by excluding very few (N = 11) participants from conducted analyses. However, unlike in research utilising data from routine care registries, participation bias could not have been avoided. The extent of this bias is difficult to assess.

As I described in Chapter 6, patient screening logs were available for only 2 of 11 participating hospital sites, indicating a considerable disparity in enrolment rates (43% for the Glasgow Royal Infirmary vs 12% for the Royal Alexandra Hospital). The differences in characteristics between the APPLE and SSNAP cohorts (as can be seen in Table 6-4) are also difficult to interpret, as the latter did not include patients with a diagnosis of TIA. Nonetheless, it seems that the APPLE sample underrepresents survivors of severe stroke.

In all decisions related to the study design and statistical analyses, I aimed to follow current recommendations as outlined in GRoLTS (441), PROBAST (95) and relevant materials on SEM (494, 495). Even so, the strengths of chosen strategies were frequently accompanied by certain drawbacks. Conducting factor analysis to obtain a latent variable representing cognitive function allowed me to handle missing data from telephone assessments, account for differing contributions of test items, and verify measurement invariance throughout follow-up.

At the same time, however, to ensure success of the procedure, I applied a data-driven approach for exclusion and collapsing of specific AMT-plus questions. The need for such adjustments suggests a certain level of redundancy across test items, and reminds that there is no evidence regarding the psychometric properties of this particular cognitive screening tool (a hybrid between the AMT-10 and a short-form MoCA). Moreover, the meaning of the individual factor scores that I derived is only relative. As such, estimates of cognitive function cannot be directly translated to scores from standardised assessment measures, including thresholds for diagnosis of cognitive impairment.

On a similar note, through accounting for differences in assessment timepoints across participants, I increased the probability of successful model convergence and good model fit (443). Unfortunately, not all preferred software output features were available for more complex models such as the one developed. This specifically relates to plotting estimated factor score means against observed means for identified trajectories. Figures included in the Results section only present observed means. Adding estimated means is recommended to aid the selection of an optimal solution from considered alternatives, differing in terms of number of classes and type of growth function, and for assessing overall model fit (441).

Further, I accounted for the possibility of non-random missingness in outcome data – an issue frequently ignored in studies involving use of LCGA. To attenuate potential bias, I initially planned to apply a method that had been tested in a previous study (453). However, emerging challenges seemed to justify introducing modifications, ultimately resulting in the use of a somewhat experimental approach. Nonetheless, compared to results of the default analysis using FIML alone, from a clinical perspective the adjusted trajectory shapes seemed more plausible, while the overall effect on the final model was limited.

Regarding additional analyses, the study sample was too small to provide sufficient power for an outcome dividing participants into four classes. While the solution to combine subgroups may have increased the likelihood of detecting significant effects, it also led to a loss of information, relevant to the uniqueness of each class. To some extent, this could have contributed to the observed lack of association between trajectory membership and measures of physical activity

and social support. Plausibly, there were also other relevant limitations, including reliance on self-report for the assessment of physical activity, and for both variables – the briefness of chosen questionnaires. Although ensuring minimal participant burden, the restricted number of incorporated items may have been insufficient to reflect interindividual variability in measured constructs.

I also experienced challenges to replicating the final model from Chapter 5. At the variable level, the effects of previous stroke and TIA could not be separated. At the path level, ensuring good model fit required specifying prior cognitive impairment as a predictor for stroke severity. As a result, it is difficult to interpret any differences in findings across the two studies. Nonetheless, the relevance of including a path from prior cognitive impairment to stroke severity is in itself an important finding, and is supported by previous research, indicating more severe strokes in patients with dementia (496, 497).

8.2.4 Future directions

As indicated in the Introduction, an immediate next step will be to repeat this analysis with inclusion of data from the 18-month assessment timepoint. This will provide an opportunity to further develop conducted procedures. Perhaps the most important planned modification relates to specifying a different type of growth function. With the exception of the low – constant improvement class, graphical representations clearly indicated that for the remaining three trajectories a single linear function could not capture observed differences in rate and/or direction of change over time. Availability of five assessment timepoints will allow to apply what seems to be a more suitable solution – modelling of piecewise trajectories (444). This involves breaking growth into specific segments. For the APPLE dataset, a better fitting model could likely be achieved through specifying one linear function for the period from baseline to the 1-month assessment, and another for the remaining duration of follow-up.

Increasing model complexity may also be advantageous in relation to three further study aspects. Firstly, I will additionally specify a model based on GMM, to enable accounting for potential within-class heterogeneity. Secondly, I will explore the use of more advanced techniques to handling missing outcome data,

allowing for explicit modelling of dropout (454). Thirdly, for variables of interest that were reassessed during follow-up (alongside cognitive performance), I will consider effects at different stages of trajectory progression rather than exclusively at baseline.

This final aspect seems of particular interest in view of identifying potential targets for intervention and, here specifically, will offer an opportunity to readdress the role of physical activity and social support (notwithstanding limitations discussed above). Two publications point to the latter as an indeed promising focus of investigation, as over a six-month period following stroke, high social support was associated with a considerable increase in functional improvement (246, 247). In one of the studies (246), the authors also found that after an initial period of recovery, participants with the lowest level of social support were likely to show a decline in functional status, beginning at around two months post-stroke. Moreover, compared to survivors of severe and moderate stroke, those with mild stroke were at highest risk of low support.

These findings appear particularly relevant in view of how trajectory shape changed after the 1-month assessment for APPLE participants in the two higher-function classes. On a broader perspective, they also point to another potential explanation for why overall better post-stroke cognitive function could be coupled with a less favourable pattern of change, that is, differences in treatment, available services, opportunities, and experiences across stroke survivors, depending on their initial presentation.

Having conducted repeated literature searches for the purpose of my thesis, I have not yet found a publication describing the use of latent growth modelling techniques for identifying trajectories of post-stroke cognitive change. This adds further merit to expanding this field of research by repeating similar studies in different datasets. Preferably, these will involve larger sample sizes and more assessment timepoints, and will offer a broad scope of investigation for modifiable determinants of outcome, the targeting of which could promote cognitive improvement and long-term preservation of high cognitive function.

8.3 Summary

This study represents a pioneer attempt to capture post-stroke cognitive changes using a latent growth modelling technique. My results speak to the heterogeneity of the investigated process. This was reflected in the identification of four cognitive trajectories, with distinct features regarding the initial severity of deficits, and the rate and pattern of changes occurring over a one-year period. In the wider context of my thesis, this research reinforces earlier conclusions regarding the complexity of the cognitive sequalae of stroke and their associations with individual characteristics. As I discuss in my final chapter, underestimating this complexity may create a gap between the potential and actual impact of prognosis research on improving post-stroke cognitive function.

Chapter 9 Discussion

Improving cognitive function is a priority objective for individuals affected by stroke (37). Prognosis research has the potential to play an important role in achieving this goal, through informing the development and implementation of appropriate interventions (96). The first pivotal steps involve gaining a better understanding of the cognitive changes that occur following stroke, and the factors that are associated with their course. In recent years, much research has been conducted to this end, with attention now turning towards applying the obtained findings to allow accurate prognosis of post-stroke cognitive outcome at an individual level, based on selected characteristics.

Through examining the existing literature with a focus on post-stroke cognitive function, I identified three under-investigated topics, where further evidence could meaningfully contribute to the foundations for this next stage of research. This included assessing the relevance of: i) potentially modifiable factors, ii) differential effects of risk factors, depending on paths of influence and co-occurrence, and iii) interindividual differences in intraindividual cognitive change over time.

9.1 Relevance of potentially modifiable factors

The unique value of identifying modifiable factors lies in their potential to serve as targets for intervention. In addressing this topic, I focused on two domains of everyday life that have received much research interest as prognostic (and possibly causal) factors for cognitive function in the general population – physical activity patterns and social engagement. In relation to the former, the associations that I observed most consistently related to the daily duration of two types of sedentary behaviour, where computer use had a positive effect on cognitive performance, and the opposite was indicated for watching TV. At the same time, I found very little evidence to support the anticipated relationship between increased physical activity and better cognitive function. As discussed in Chapter 4, based on these findings I reached two main conclusions: 1) it is possible for sedentary behaviour to be associated with post-stroke cognitive function independently of physical activity, and 2) sedentary behaviour may

have opposing effects on cognitive function depending on whether it is mentally passive or active.

In relation to social engagement, from a number of measures representing both ends on the objective-subjective continuum, I found that cognitive task performance was most consistently associated with loneliness. Importantly, feeling lonely was reported by one in four stroke survivors in the study sample – a concerning observation given the distressing nature of the experience itself, as well as previous evidence suggesting its detrimental relationship with not just cognition (301-303), but also many other individual outcomes, such as depression, diminished immunity, cardiovascular disease, and mortality (297, 298, 300, 498).

For the interpretation of my results from both studies, it is important to note that the associations I observed between cognitive performance and predictors of interest overall were weak, and there was some heterogeneity in findings across different tasks. Moreover, as discussed in sections 4.3.2.3 and 4.3.4.3, both studies had considerable methodological limitations, many of which related to use of data from a general-purpose cohort (UK Biobank). One, involving the use of suboptimal measures of physical activity and social support, also posed an issue in the APPLE study. Here, this specifically related to the briefness of both chosen questionnaires, and for the assessment of physical activity – reliance on self-report. Coupled with a relatively small sample size, this may explain my neutral results, where neither of these factors predicted the course of post-stroke cognitive change over a one-year period.

In view of this, my findings from the UK Biobank analyses on the associations between habitual physical activity, sedentary behaviour, and social engagement on post-stroke cognitive function cannot be considered as definitive. However, they provide grounds for instigating further focussed studies in this area. Investigating the impact of breaking up mentally passive sedentary behaviour with cognitive and light-intensity physical activity seems of particular interest, as a more feasible and sustainable alternative strategy to structured, supervised exercise sessions for stroke survivors with greater activity limitations and participation restrictions. Promoting social support, on the other hand, may be considered as an additional interventional component, having the potential to

improve adherence to treatment programmes (499-501), alongside a possible, more direct effect on cognitive function.

9.2 Differential effects of risk factors

The potential for risk modification was also a key consideration that motivated me to focus on the cognitive consequences of cardiovascular diseases. While there is a strong premise to assume their relevance to post-stroke cognitive function, existing evidence on this topic appeared inconclusive or conflicting (23, 145). I recognised that this may be due to the complex nature of these associations, with some effects on cognition preceding index stroke, interactions between diseases, and multiple paths of influence (direct and indirect). In relation to the latter, I further assumed that the effects of cardiovascular risk factors may not be unequivocally detrimental, as pathological processes can trigger endogenous adaptive mechanisms; for example the theory around TIA inducing ischaemic tolerance (343), and large vessel cerebrovascular disease leading to the development of collateral circulation (347, 348). This inspired me to develop a model, allowing me to test the indirect effects of cardiovascular risk factors on acute post-stroke cognitive function, as mediated through stroke severity and a history of dementia, and to assess moderation effects due to comorbidity.

As I describe in Chapter 5, my results indicated that poorer cognitive performance was associated with: atrial fibrillation through increased stroke severity, previous stroke through an increased risk of prevalent dementia, as well as age through both mediators. Importantly, my findings also supported the hypothesis that some indirect effects of cardiovascular risk factors may be favourable. Specifically, vascular disease in the presence of hypertension and absence of diabetes, as well as previous TIA seemed to be associated with better acute cognitive performance through reduced stroke severity.

I conducted a similar analysis in the APPLE dataset, which I described in Chapters 7 and 8, using an alternative outcome – the pattern of post-stroke cognitive change over one year, dichotomised to reflect an overall higher and overall lower level of cognitive function. In addition to confirming the relevance of stroke severity and prior cognitive impairment, the main comparable finding

related to an indirect association between atrial fibrillation and lower overall cognitive function, through increased stroke severity. Conversely, one of my most unique observations concerned the possible opposing effects of age – associated with higher cognitive function through reduced stroke severity, and with lower function through an increased risk of prior cognitive impairment. To add further complexity, prior cognitive impairment was in turn linked to greater stroke severity.

The differences in findings across the two studies are difficult to account for, as I was unable to precisely replicate the first model in the APPLE dataset. This was only one of many challenges associated with conducting these analyses and the interpretation of results. Another significant limitation related to uncertainty in the measurement of variables of interest, particularly regarding pre-stroke cognitive function. Similar to most predictors, coding of this factor relied on the presence of information in medical records, while evidence suggests that underdiagnosis and/or under-recording of cognitive impairment, even in its severe form, is a common issue (389).

Moreover, I did not have access to data relevant to verifying my initial assumptions, regarding the possible impact of previous TIA and vascular disease on alleviating stroke severity, and thus better acute cognitive performance. While I developed my hypotheses based on an assumed role of endogenous protective adaptations, treatment effects (e.g. from prescribed medication) provide a highly plausible alternative explanation for my findings (386, 387).

Overall, the results of my analyses still leave much uncertainty. Nonetheless, they are not without research implications. My findings suggest that commonly applied basic multivariable models constitute an overly reductionist approach to investigating the associations between cardiovascular diseases and post-stroke cognitive function. Consequently, the importance of the former may be easily underestimated. This conclusion is also likely to apply to assessing the relevance of lifestyle factors, as their effect on cognition is plausibly (at least in part) exerted through modifying cardiovascular risk.

In view of this, I would recommend future application of moderated mediation analyses for studying the associations between modifiable factors and post-

stroke cognitive function, considering multiple putative paths of influence and interaction effects. To gain a better understanding of the mechanisms that underly investigated relationships, it moreover seems necessary to account for neuroimaging evidence of structural, metabolic, and/or functional brain abnormalities. Inclusion of such variables in statistical models as mediators could contribute to explaining how modifiable factors affect post-stroke cognition through their impact on progression of neuropathological processes.

My findings regarding the complexity of between-predictor and predictor-outcome relationships may also have clinical implications. Assuming observed associations reflected true effects, stroke survivors with a history of TIA and vascular disease with hypertension could be considered as having a relatively low risk of future post-stroke cognitive disorders on account of less severe index strokes. However, the presence of these conditions is likely associated with progressive neurological damage, occurring long-term. The conclusion that for specific subgroups of stroke survivors the actual risk of cognitive deterioration may be underappreciated was also implied by my findings from the APPLE study.

9.3 Trajectories of post-stroke cognitive change

As I emphasised in Chapter 2 in the section on stratified medicine, while sharing certain characteristics, clinically defined populations are in many ways heterogeneous. This can manifest in variability in the clinical course of a condition or - expanding this notion beyond the context of disease – in the change that occurs in individual status over time. In relation to cognitive function following stroke, it is moreover apparent that in addition to differences between individuals, there is also heterogeneity in the pattern of change on a within-individual level. Latent growth modelling techniques were developed to capture these aspects – the interindividual differences in intraindividual change over time (440). Yet, it appears they have not been previously employed to investigate the topic of post-stroke cognition.

Through applying LCGA in the APPLE dataset, I identified four distinct trajectory classes reflecting cognitive change during the first year following stroke: i) a high – declining class, with high cognitive function soon after stroke, improving over following weeks, and thereafter declining; ii) a mid-high – stable class, with

some cognitive problems soon after stroke, followed by a period of improvement, after which cognitive function remained relatively stable; iii) a mid-low – slowing improvement class, with comparatively poor initial cognitive function, which after a stage of steeper improvement, continued to improve at a slower rate; and iv) a low – constant improvement class, with severe cognitive impairment soon after stroke, followed by improvement in cognitive function at a relatively constant rate. During the follow-up period, the order of classes according to level of cognitive performance prevailed, meaning that at no point did a class with poorer initial cognitive function exceed in performance a class with higher initial function. In general, study participants representing the two classes with lower overall cognitive function were characterised by older age, lower education, greater stroke severity, and pre-stroke cognitive impairment.

Given the pilot nature of this research, similar analyses need to firstly be conducted in the full APPLE dataset (including an 18-month assessment), followed by testing in independent participant samples, before confidence in my findings will seem justifiable. Nonetheless, I consider that the particular merit of this study does not lie in closing, but rather opening questions. These apply to some of the basic, widespread assumptions made in prognosis research into post-stroke cognition: *Is the absence of diagnosed cognitive impairment at a specific point in time equivalent to a good outcome? Is prognostic accuracy achievable assuming direct, unconditional predictor-outcome associations? How applicable are average estimates to any individual, given such heterogeneity within a stroke population?* These questions, together with my initial conclusions from reviewing the existing literature, and findings from my earlier studies, evoked reflections on the role and direction of development for this area of research.

9.4 Appraisal of prognosis research into post-stroke cognition: a pause before deciding on future directions

Viewing my thesis as a whole, it may seem that the existing research work I described across the first three chapters called for a different direction of investigations. In accordance with the PROGRESS framework (96), which was the focus of Chapter 2, I inferred that prognosis research into post-stroke cognition has presently reached the beginning of the third theme, related to the

development of prognostic models. Conversely, the three topics I chose to address in my studies can be classed as pertaining to the two preceding themes of fundamental prognosis research and prognostic factor research. This gives rise to an unavoidable question: why look back instead of moving forward?

The reason for this was that I considered the next obvious steps – external validation and potential updating of existing prognostic rules – premature, with two crucial issues remaining open to debate:

1. is accurate prognosis of individual post-stroke cognitive outcomes viable?

2. is the ability to predict individual post-stroke cognitive outcomes valuable?

As I found, the results of my studies only added fuel to these uncertainties, particularly in relation to viability.

9.4.1 Examining the viability of individual outcome prognosis

9.4.1.1 Post-stroke cognitive impairment

As I described in Chapter 3, to date only one prognostic rule for predicting post-stroke cognitive impairment has been externally validated – the CHANGE score (164). It presented only fair discriminatory power (AUROC of 0.75), despite the prediction being relatively short-term, with the outcome assessed between three to six months post-stroke. While this is the result of only one study, and comparable data are unavailable for other identified prognostic rules, this observation is not singular in suggesting limited potential for such tools to have good accuracy (AUROC of at least 0.80). Further support is gained from a systematic review of 61 prognostic models for predicting dementia in the general population, where good performance in an independent dataset was rarely reported (152), followed by several deductive arguments.

One issue has been repeatedly emphasised across all my studies – the nature of the matter is extremely complex. Cognitive function is shaped by multiple interlinked experiences and exposures, with influences present throughout the entire lifespan (502). Even when just focusing on selected prognostic factors,

appearing most relevant to post-stroke cognition either based on temporal proximity or putative underlying causal mechanisms, modelling their effects is a challenging task. As I demonstrated, this relates to the possibility of differential associations for a given predictor with post-stroke cognition, stemming from varying paths of influence, between-variable interactions, and the heterogeneity of the population.

In stark contrast to this complexity, is the relative simplicity of prognostic rules. Indeed, maximising simplicity is an intentionally sought property in the design of prognostic rules to ensure feasibility of use. Arguably, for the purpose of outcome prognosis, where the effect of a variable is indirect, conditional, or non-universal, it can be recognised as having poor prognostic value, and thus to not merit attention. This approach does not seem unreasonable, given that there are alternative predictors, such as age, stroke severity, and cerebral atrophy, which have a strong, consistently proven association with post-stroke cognition.

Nonetheless, to some extent, relativity of effects will also apply to such factors, and while ignoring this issue will not negate their prognostic utility, it may diminish it. For example, it is worth considering whether performance of the CHANGE score could be improved by incorporating interaction effects between some of the predictors it already includes, namely, between education and both white matter hyperintensities and global cortical atrophy. Indeed, many studies in the field of cognitive reserve have indicated that education attenuates the effects of neuropathological changes on cognitive function (503, 504)

Alongside a reductionist approach to modelling the effects of predictors, perhaps the greatest commonly applied simplification relates to capturing the outcome, with post-stroke cognition forced to fit the dimensions of a single, binary event. Importantly, however, estimating that an individual is at low risk of post-stroke cognitive impairment at a specific point in time gives no assurance that they are at low risk of cognitive decline. Findings from my latent class growth analysis support the notion of such duality, indicating declining performance in individuals with fewest cognitive deficits, who at the same time were younger, more educated, had fewer comorbidities, and had less severe index strokes than participants representing classes with overall lower, albeit improving function.

Overlooking the direction of cognitive change over time is also likely to entail a certain short-sightedness of predictions. Despite decline in cognitive function, a prognostic rule may correctly predict that according to standard criteria an individual will be classed as cognitively intact at six months; but will this still apply at one year if changes continue along the same trajectory?

9.4.1.2 Post-stroke delirium

The above concerns seem of lesser relevance to prognosis of delirium in acute stroke. Intrinsically, the aim of a prognostic rule in this context is to predict an outcome typically occurring within a matter of days. The sudden onset and transient nature of delirium is also more akin to the concept of a single event. For similar reasons, there is merit in prioritising factors that are affecting the individual in the acute phase of stroke, such as existing comorbidities, sensory impairments, infection, or dehydration, rather than life-course influences.

Nonetheless, external validations of prognostic rules for predicting delirium in older adult inpatients indicate that in most cases performance is also only fair (153). This seems to remind us that as in relation to mild cognitive impairment and dementia, the development, progression and determining factors of delirium are not yet fully understood. Whether accurate prognosis of individual outcomes is possible without such an understanding is debatable.

9.4.2 Examining the value of individual outcome prognosis

The issues of viability and value of applying prognostic rules to predict individual post-stroke cognitive outcome are intertwined. At a minimum, produced estimates can only be considered of value if they are better than chance. Conversely, even excellent discrimination and calibration cannot ensure that an estimated risk is worth knowing. Assuming that satisfactory prognostic accuracy is achievable, below I consider the potential importance and implications of applying prognostic rules from individual, research, and policy perspectives.

9.4.2.1 A clinical and personal perspective

Individual prognosis of post-stroke cognitive impairment

A qualitative study investigating the views of stroke survivors, carers and clinicians on use of prognostic tools for predicting post-stroke dementia in routine practice highlighted a shared, paramount concern (175): is knowledge of the risk worth the entailed psychological burden, given there is currently no established treatment to change the outcome? Familiarity with the seven existing prognostic rules addressing post-stroke cognitive impairment seems to only reinforce a sense of inevitability.

As I presented in the Results section of Chapter 3, nearly all the considered predictors were non-modifiable. Moreover, they relied on information collected around the time of index stroke. This is naturally understandable, and even ideal, in view of an intention to apply a prognostic rule in the acute phase. However, confidence in such predictions would also imply that regardless of what an individual does after a stroke (e.g. close adherence to prescribed medications or lifestyle changes), this will not have a meaningful impact on their risk of cognitive impairment. From this perspective, contentiously, it seems we should hope that existing rules would not perform well in external validation.

Appropriate initiation of preventive interventions, however, is not the only aim of estimating the individual probability of an unfavourable outcome. In the same qualitative study (175), participants expressed that awareness of the risk of dementia could help in planning for the future and finding useful coping strategies. This view does not yet seem to explicitly support the use of prognostic tools, at least in their present form.

Two of the existing rules aimed to predict cognitive impairment up to one year following stroke, while the five remaining ones focused on the outcome occurring after three to six months. Given this relatively short time span, it is questionable whether a prognostic rule is indeed useful in alerting a stroke survivor and/or carer to the prospect of severe cognitive difficulties, and the consequent need for preparation. Plausibly, many of the deficits informing the diagnosis of dementia, e.g. at 6 months, would have already been present in the

acute phase of stroke. Assuming these impairments were recognised through routine cognitive screening or personal observations, practical adjustments to daily lives of the stroke survivor and their family/carers could have been made regardless of the prognosis.

Another issue relates to whether it is ever truly appropriate to assume that deterioration of cognitive function is of no concern to a stroke survivor, even if the risk estimated by a prognostic rule is low. A recent study reported that one year following a severe stroke, the risk of dementia is a staggering 50 times higher than in the age and sex-matched population (20). In this context, however, it is particularly noteworthy that the incidence of dementia was found to be comparatively 3.5 times higher just after TIA, increasing to approximately 6 times for minor stroke (NIHSS score < 3). Further supporting this argument are my findings regarding trajectories of cognitive change, presented in Chapter 8. Improvement was coupled with relatively poor cognitive function, while high function – with decline. As such, it is questionable whether representing any subgroup constitutes a satisfying outcome, with the answer likely depending on individual point of view.

Individual prognosis of post-stroke delirium

As in the case of viability, the evaluation of prognostic rules for predicting delirium takes a somewhat different course. On one hand, there are recommended interventions for delirium (64). On the other, recognition of this condition in the context of stroke can often be difficult (147, 505). Identifying high risk individuals could encourage closer and more frequent monitoring of behaviour and cognitive function, reducing the likelihood of delirium being missed, and thus remaining untreated. Yet, perhaps even more importantly, risk awareness could contribute to delirium prevention.

A Cochrane review (506) concluded there was strong evidence supporting multi-component interventions to prevent delirium in general, geriatric, and surgical ward settings. While the specific content of considered interventions varied, some elements were shared across multiple studies, including: provision of individualised care, reorientation (repeatedly informing a patient about the environment/circumstances they are in), avoidance of sensory deprivation,

maintaining appropriate nutrition and hydration, mobilisation, and promoting sleep hygiene. Since the publication of the review, one study (N = 108) has further tested a multi-component intervention specifically in an acute stroke setting, indicating a reduction in the incidence rate of delirium by 16.7% (507).

However, despite the possibility of preventing or treating post-stroke delirium, there still seems to be insufficient justification for estimating the individual risk of its development in clinical practice. Looking at the intervention components listed above, it can be argued that all acute stroke patients should be entitled to such care; particularly, as the absolute risk of delirium in this group is high, being at approximately 25% (21).

9.4.2.2 A research perspective

The value of applying prognostic rules in a research context is typically indicated in relation to clinical trials. Due to heterogeneity within a clinical population, the average benefit of a tested intervention reported for a study sample may in many cases be a poor representation of the likely treatment effect in a particular individual (508). Exploring differences in treatment effects by employing conventional subgroup analysis, where participants are repeatedly categorised based on one variable at a time (e.g. sex or history of hypertension), is recognised as having significant limitations (509). Specifically, an increased risk of false positive findings due to multiple testing is coupled with a reduced ability to detect real heterogeneity in treatment effects (510). The latter results from any single variable likely having only a small influence on treatment effect, as well as from possible similarities across analysed subgroups regarding many other pertinent characteristics.

By simultaneously accounting for multiple factors, prognostic rules overcome these limitations. Importantly, the single property they capture - the risk of an unfavourable outcome – is considered highly relevant to explaining variation in treatment effect (509). As I described in Chapter 2 in the section on stratified medicine, where the relative effect of a treatment is found to be similar across individuals, those with the highest initial risk of an unfavourable outcome will experience the greatest absolute benefit. This is a key consideration for clinical decision-making. Given that most interventions are associated with some

adverse effects, for a person whose probability of an unfavourable outcome is low, the estimated benefit from the treatment may be too low to outweigh the potential risks.

The relevance of the individual risk of a future outcome to treatment effect is recognised as practically universal (509). As follows, prognostic rules for predicting cognitive disorders may indeed contribute to clinical trials of interventions targeting cognitive function following stroke. The extent of their application, however, may depend on intervention content. Arguably, the potential impact of some treatments may by limited by the presence of factors that contribute to the overall increased probability of a future cognitive disorder, thus disrupting the expected association between a higher risk and greater absolute benefit. To illustrate this point, I describe an example based on prevention - currently considered a priority approach to reducing the high global burden of dementia (511, 512).

Similarly as in the case of delirium, recognition of the multifactorial aetiology of dementia has encouraged the development of multi-component interventions for risk reduction (512, 513). Focusing on such aspects as physical activity, diet, control of cardiovascular risk factors, and cognitive training, these interventions heavily rely on self-management, with study participants required to individually implement changes in their everyday life (514-517). Following stroke, however, individuals with initial cognitive problems may experience challenges to fully participating in interventions based on self-management (518, 519). As follows, despite likely being at greater risk of future dementia, they may benefit from such strategies less than stroke survivors for whom there are no cognitive barriers to implementing and maintaining lifestyle recommendations.

Recognising that there are distinct subpopulations among stroke survivors, further draws attention to the possibility of having to address differing intervention aims. Reflecting on the trajectories of cognitive change that I identified in Chapter 8, for individuals with overall high cognitive function, in the sub-acute stage of stroke it could be a priority to maintain the current level of performance and avoid cognitive decline. For individuals representing an overall low cognitive function trajectory, however, it seems it would be

important to focus on improving performance. Whether one type of intervention could address both aims with equal success is open to question.

9.4.2.3 A healthcare policy perspective

An increasing burden of cognitive disorders has major implications for health and social services (520, 521). The current prioritisation of prevention for strategic policy-making has been coupled with a desire to identify high-risk individuals for targeting interventions (511, 512, 522). Interestingly, in “From evidence into action: opportunities to protect and improve the nation’s health” released in 2014 by Public Health England, one of the stated aims in relation to dementia prevention was to develop a personalised risk assessment calculator for incorporation into the NHS Health Check (523). However, a rationale for this decision had not been described.

In general, health policies aimed at dementia prevention focus on improving public and professional awareness of the disorder, and promoting behaviours that can lead to the reduction of modifiable risk factors, such as physical activity, a healthy diet, smoking cessation, and limited alcohol consumption (511). Importantly, such recommendations appear to be universally beneficial, and to entail a low risk of adverse effects. Moreover, the relative risk of cognitive impairment and decline for stroke survivors is already high compared to the general population. Therefore, a similar question arises as for implementation of multicomponent interventions for preventing delirium – why should such strategies be targeted only at a specific subgroup of stroke survivors, who are at highest risk of dementia?

This is not to say that prognosis research has not and will not be crucial to the development of effective healthcare policies for the prevention of post-stroke cognitive disorders. Based on indications from the PROGRESS series, key contributions may include (96, 99): i) estimating average prognosis to model the population burden of cognitive disorders and associated service requirements, ii) comparisons of different healthcare systems to inform how variation in standard care influences cognitive outcomes, iii) identifying modifiable risk factors to target in preventive, policy-level interventions, and iv) assessing the impact of such interventions.

9.5 Conclusions

To date, there is no specific, recommended intervention to improve cognitive function following stroke. A seemingly clear path through which prognosis research may contribute to changing this reality involves identifying modifiable prognostic factors, causally associated with post-stroke cognition, to serve as treatment targets. In this context, findings from my UK Biobank studies provided preliminary evidence to support further investigation of three factors – mentally passive sedentary behaviour and loneliness as associated with poorer cognitive function, and mentally active sedentary behaviour as associated with better function.

These observations already highlight one important challenge to understanding predictor-outcome relationships – what could be considered a single variable, such as the level of habitual physical activity or social engagement, may have a number of different components, each characterised by distinct properties. Without recognising their unique relevance, true and important associations could be easily missed, misrepresented or misinterpreted.

It is yet perhaps the next step that seems to be the hardest based on previous research endeavours – translating findings from observational studies (as the ones I conducted) to successful interventions, improving individual outcomes in a clinically meaningful way. The results of my subsequent studies seem to provide insight on why this gap is so difficult to bridge, as well as suggestions for how this issue could be addressed.

Firstly, it is important to consider how and when a potentially causal factor can affect post-stroke cognitive outcome. As I found through my moderated mediation analyses, a lack of evidence-based assumptions regarding these aspects to inform the development of a statistical model may also lead to misidentification of relevant associations. This shortcoming is then likely to have implications at the stage of intervention design, with accurate answers to “how” and “when” crucial to ensuring applicability of a considered approach.

For example, there has been much controversy regarding active blood pressure lowering in acute stroke, with studies reporting mixed results on the impact of

this strategy on early and long-term outcomes (including, e.g. neurological deterioration, death, and dependency) (524, 525). Findings from a recent study, which I described in Chapter 5, indicated a possible explanation for such inconsistencies - the impact of blood pressure on clinical outcomes is highly dependent on reperfusion (380). With a favourable effect of higher blood pressure in patients with reperfusion, and the opposite observed for patients without it, the authors suggested that active blood-pressure lowering should not be considered prior to reperfusion treatment.

Secondly, stroke survivors constitute a heterogenous population, which can manifest in differential patterns of cognitive change over time, and varying relevance of specific prognostic factors. This observation undermines the possibility of developing an effective one-size-fits-all intervention. While risk-stratification based on prognostic rule estimates may account for some of the variation in treatment effects, it seems that a more holistic and in-depth understanding of individual profiles may be necessary to appropriately tailor future interventions (526). In addition to factors that I addressed in my thesis, neuroimaging findings may play an important role in gaining this understanding, allowing to account for interindividual differences in the extent of neuropathological changes, their type, including whether they are reversible (e.g. metabolic abnormalities) or irreversible, and their manifestation.

In summary, I believe that in the present context of prognosis research in post-stroke cognition, fundamental and prognostic factor investigations still have a priority role to play in advancing the search for ways to improve individual outcomes. For contributions to be meaningful, however, it seems important to reconsider our approach to capturing cognitive changes following stroke and their associations with individual characteristics. Overlooking information that speaks to the complexity of the subject matter can significantly limit the real-world application of research findings. It is then, when the practical advantages of conceptual simplifications come at too high a cost.

Appendices

Appendix 1: Chapter 3, PRISMA checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	52
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	N/A
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	52, 53
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	53
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	53
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	54, 55
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	54
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	218-224 (Appendix 2)
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	54, 55

Section/topic	#	Checklist item	Reported on page #
METHODS			
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	53, 55; RP
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	55, 56
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	56
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	56
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	N/A
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	N/A
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	57, 58
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	59-69
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	60, 61, 70
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	66-68
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	N/A

Section/topic	#	Checklist item	Reported on page #
RESULTS			
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	N/A
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	71-74
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	72-74
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	74, 75
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	N/A

RP indicates Review Protocol (https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020170428).

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

Appendix 2: Chapter 3, search strategy

Medline (via OVID) search strategy

1. cerebrovascular disorders/ OR exp basal ganglia cerebrovascular disease/ OR exp brain ischemia/ OR exp intracranial arterial diseases/ OR exp "intracranial embolism and thrombosis"/ OR exp intracranial hemorrhages/ OR stroke/ OR exp brain infarction/ OR vasospasm, intracranial/
2. (stroke OR post?stroke OR cerebrovasc$ OR brain vasc$ OR cerebral vasc$ OR cva$ OR apoplex$ OR SAH).ti,ab.
3. ((brain$ OR cerebr$ OR cerebell$ OR intracran$ OR intracerebral) adj5 (isch?emi$ OR infarct$ OR thrombo$ OR emboli$ OR occlus$)).ti,ab.
4. ((brain$ OR cerebr$ OR cerebell$ OR intracerebral OR intracranial OR subarachnoid) adj5 (h?emorrhage$ OR h?ematoma$ OR bleed$)).ti,ab.
5. ((transi$ adj3 isch?em$ adj3 attack$) OR TIA$1).ti,ab.
6. 1 OR 2 OR 3 OR 4 OR 5
7. ((validat$ OR predict$ OR prognos$ OR rule$) adj3 (outcome$ OR risk$ OR model$)).ti,ab.
8. (prognos$ AND (method$ OR history OR variable$ OR criteria OR scor$ OR characteristic$ OR finding$ OR factor$ OR model$)).ti,ab.
9. ((history OR variable$ OR criteria OR scor$ OR characteristic$ OR finding$ OR factor$) adj3 (predict$ OR model$ OR decision$ OR identif$ OR prognos$)).ti,ab.
10. (decision$ adj3 (model$ OR clinical$)).ti,ab.
11. (stratification OR discriminat$ OR calibration).ti,ab.
12. ROC curve/
13. (c-statistic OR c statistic OR area under the curve OR AUC).ti,ab.
14. (indices OR algorithm OR multivariable).ti,ab.
15. 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14
16. exp dementia/
17. delirium/
18. delirium, dementia, amnestic, cognitive disorders/
19. exp cognition disorders/
20. exp cognition/
21. memory/

22. dement$.ti,ab.
23. (Alzheimer$ OR AD).ti,ab.
24. deliri$.ti,ab.
25. ((cognit$ OR memory OR mental OR brain) adj3 (func$ OR perform$ OR abilit$ OR declin$ OR reduc$ OR impair$ OR disorder$ OR fail$ OR los$ OR deficit$ OR stop$ OR progress$ OR improve$)).ti,ab.
26. mental perform$.ti,ab.
27. (memory adj3 (complain$ or declin$ or function$)).ti,ab.
28. 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27
29. 6 AND 15 AND 28

Embase (via OVID) search strategy

1. cerebrovascular disease/ OR exp basal ganglion hemorrhage/ OR exp brain hematoma/ OR exp brain hemorrhage/ OR exp brain infarction/ OR exp brain ischemia/ OR cerebral artery disease/ OR exp cerebrovascular accident/ OR exp occlusive cerebrovascular disease/ OR vertebrobasilar insufficiency/ OR stroke/ OR stroke patient/ OR stroke unit/
2. (stroke OR post?stroke OR cerebrovasc$ OR brain vasc$ OR cerebral vasc$ OR cva$ OR apoplex$ OR SAH).ti,ab.
3. ((brain$ OR cerebr$ OR cerebell$ OR intracran$ OR intracerebral) adj5 (isch?emi$ OR infarct$ OR thrombo$ OR emboli$ OR occlus$)).ti,ab.
4. ((brain$ OR cerebr$ OR cerebell$ OR intracerebral OR intracranial OR subarachnoid) adj5 (h?emorrhage$ OR h?ematoma$ OR bleed$)).ti,ab.
5. ((transi$ adj3 isch?em$ adj3 attack$) OR TIA$1).ti,ab.
6. 1 OR 2 OR 3 OR 4 OR 5
7. ((validat$ OR predict$ OR prognos$ OR rule$) adj3 (outcome$ OR risk$ OR model$)).ti,ab.
8. (prognos$ AND (method$ OR history OR variable$ OR criteria OR scor$ OR characteristic$ OR finding$ OR factor$ OR model$)).ti,ab.
9. ((history OR variable$ OR criteria OR scor$ OR characteristic$ OR finding$ OR factor$) adj3 (predict$ OR model$ OR decision$ OR identif$ OR prognos$)).ti,ab.
10. (decision$ adj3 (model$ OR clinical$)).ti,ab.
11. (stratification OR discriminat$ OR calibration).ti,ab.
12. receiver operating characteristic/

13. (c-statistic OR c statistic OR area under the curve OR AUC).ti,ab.
14. (indices OR algorithm OR multivariable).ti,ab.
15. 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14
16. exp dementia/
17. delirium/
18. exp cognitive defect/
19. exp cognition/
20. memory/
21. dement$.ti,ab.
22. (Alzheimer$ OR AD).ti,ab.
23. deliri$.ti,ab.
24. ((cognit$ OR memory OR mental OR brain) adj3 (func$ OR perform$ OR abilit$ OR declin$ OR reduc$ OR impair$ OR disorder$ OR fail$ OR los$ OR deficit$ OR stop$ OR progress$ OR improve$)).ti,ab.
25. mental perform$.ti,ab.
26. (memory adj3 (complain$ or declin$ or function$)).ti,ab.
27. 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26
28. 6 AND 15 AND 27

PsycINFO (via EBSCO) search strategy

S1 DE "Cerebrovascular Disorders" OR DE "Cerebral Arteriosclerosis" OR DE "Cerebral Hemorrhage" OR DE "Cerebral Ischemia" OR DE "Cerebrovascular Accidents" OR DE "Subarachnoid Hemorrhage"

S2 TI (stroke OR post#stroke OR cerebrovasc* OR "brain vasc*" OR "cerebral vasc*" OR cva* OR apoplexy OR SAH) OR AB (stroke OR post#stroke OR cerebrovasc* OR "brain vasc*" OR "cerebral vasc*" OR cva* OR apoplexy OR SAH)

S3 TI ((brain* OR cerebr* OR cerebell* OR intracran* OR intracerebral) N5 (isch#emi$ OR infarct* OR thrombo* OR emboli* OR occlus*)) OR AB ((brain* OR cerebr* OR cerebell* OR intracran* OR intracerebral) N5 (isch#emi$ OR infarct* OR thrombo* OR emboli* OR occlus*))

S4 TI ((brain* OR cerebr* OR cerebell* OR intracerebral OR intracranial OR subarachnoid) N5 (h#emorrhage* OR h#ematoma* OR bleed*)) OR AB

((brain* OR cerebr* OR cerebell* OR intracerebral OR intracranial OR subarachnoid) N5 (h#emorrhage* OR h#ematoma* OR bleed*))

S5 TI ((transi* N3 isch#em* N3 attack*) OR TIA) OR AB ((transi* N3 isch#em* N3 attack*) OR TIA)

S6 S1 OR S2 OR S3 OR S4 OR S5

S7 TI ((validat* OR predict* OR prognos* OR rule*) N3 (outcome* OR risk* OR model*)) OR AB ((validat* OR predict* OR prognos* OR rule*) N3 (outcome* OR risk* OR model*))

S8 TI (prognos* AND (method* OR history OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor* OR model*)) OR AB (prognos* AND (method* OR history OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor* OR model*))

S9 TI ((history OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor*) N3 (predict* OR model* OR decision* OR identif* OR prognos*)) OR AB ((history OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor*) N3 (predict* OR model* OR decision* OR identif* OR prognos*))

S10 TI (decision* N3 (model* OR clinical*)) OR AB (decision* N3 (model* OR clinical*))

S11 TI (stratification OR discriminat* OR calibration) OR AB (stratification OR discriminat* OR calibration)

S12 TI ("c-statistic" OR "c statistic" OR "area under the curve" OR AUC) OR AB ("c-statistic" OR "c statistic" OR "area under the curve" OR AUC)

S13 TI (indices OR algorithm OR multivariable) OR AB (indices OR algorithm OR multivariable)

S14 S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13

S15 DE "Dementia" OR DE "Presenile Dementia" OR DE "Pseudodementia" OR DE "Semantic Dementia" OR DE "Senile Dementia" OR DE "Vascular Dementia"

S16 DE "Neurocognitive Disorders" OR DE "Delirium" OR DE "Memory Disorders" OR DE "Cognitive Impairment"

S17 DE "Memory" OR DE "Memory Decay"

S18 DE "Cognition"

S19 TI dement* OR AB dement*

S20 TI (alzheimer* OR AD) OR AB (alzheimer* OR AD)

S21 TI deliri* OR AB deliri*

S22 TI (((cognit* OR memory OR mental OR brain) N3 (func* OR perform* OR ability* OR declin* OR reduc* OR impair* OR disorder* OR fail* OR los* OR deficit* OR stop* OR progress* OR improve*))) OR AB (((cognit* OR memory OR mental OR brain) N3 (func* OR perform* OR ability* OR declin* OR reduc* OR impair* OR disorder* OR fail* OR los* OR deficit* OR stop* OR progress* OR improve*)))

S23 TI "mental perform*" OR AB "mental perform*"

S24 TI ((memory N3 (complain* or declin* or function*))) OR AB ((memory N3 (complain* or declin* or function*)))

S25 S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR 24

S26 S6 AND S14 AND S25

CINAHL (via EBSCO) search strategy

S1 (MH "Cerebrovascular Disorders") OR (MH "Basal Ganglia Cerebrovascular Disease+") OR (MH "Cerebral Ischemia+") OR (MH "Intracranial Arterial Diseases+") OR (MH “Intracranial Embolism and Thrombosis+”) OR (MH “Intracranial Hemorrhage+") OR (MH “Stroke+) OR (MH “Cerebral Vasospasm”)

S2 TI (stroke OR post#stroke OR cerebrovasc* OR "brain vasc*" OR "cerebral vasc*" OR cva* OR apoplexy OR SAH) OR AB (stroke OR post#stroke OR cerebrovasc* OR "brain vasc*" OR "cerebral vasc*" OR cva* OR apoplexy OR SAH)

S3 TI ((brain* OR cerebr* OR cerebell* OR intracran* OR intracerebral) N5 (isch#emi$ OR infarct* OR thrombo* OR emboli* OR occlus*)) OR AB ((brain* OR cerebr* OR cerebell* OR intracran* OR intracerebral) N5 (isch#emi$ OR infarct* OR thrombo* OR emboli* OR occlus*))

S4 TI ((brain* OR cerebr* OR cerebell* OR intracerebral OR intracranial OR subarachnoid) N5 (h#emorrhage* OR h#ematoma* OR bleed*)) OR AB

((brain* OR cerebr* OR cerebell* OR intracerebral OR intracranial OR subarachnoid) N5 (h#emorrhage* OR h#ematoma* OR bleed*))

S5 TI ((transi* N3 isch#em* N3 attack*) OR TIA) OR AB ((transi* N3 isch#em* N3 attack*) OR TIA)

S6 S1 OR S2 OR S3 OR S4 OR S5

S7 TI ((validat* OR predict* OR prognos* OR rule*) N3 (outcome* OR risk* OR model*)) OR AB ((validat* OR predict* OR prognos* OR rule*) N3 (outcome* OR risk* OR model*))

S8 TI (prognos* AND (method* OR history OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor* OR model*)) OR AB (prognos* AND (method* OR history OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor* OR model*))

S9 TI ((history OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor*) N3 (predict* OR model* OR decision* OR identif* OR prognos*)) OR AB ((history OR variable* OR criteria OR scor* OR characteristic* OR finding* OR factor*) N3 (predict* OR model* OR decision* OR identif* OR prognos*))

S10 TI (decision* N3 (model* OR clinical*)) OR AB (decision* N3 (model* OR clinical*))

S11 TI (stratification OR discriminat* OR calibration) OR AB (stratification OR discriminat* OR calibration)

S12 (MH “ROC Curve”)

S13 TI (("c-statistic" OR "c statistic" OR "area under the curve" OR AUC)) OR AB (("c-statistic" OR "c statistic" OR "area under the curve" OR AUC))

S14 TI (indices OR algorithm OR multivariable) OR AB (indices OR algorithm OR multivariable)

S15 S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14

S16 (MH "Dementia+")

S17 (MH "Delirium")

S18 (MH “Delirium, Dementia, Amnestic, Cognitive Disorders”)

S19 (MH “Cognition Disorders+”)

S20 (MH “Cognition”)

S21 (MH “Memory”) OR (MH “Memory Disorders”)

S22 TI dement* OR AB dement*

S23 TI (alzheimer* OR AD) OR AB (alzheimer* OR AD)

S24 TI deliri* OR AB deliri*

S25 TI (((cognit* OR memory OR mental OR brain) N3 (func* OR perform* OR ability* OR declin* OR reduc* OR impair* OR disorder* OR fail* OR los* OR deficit* OR stop* OR progress* OR improve*))) OR AB (((cognit* OR memory OR mental OR brain) N3 (func* OR perform* OR ability* OR declin* OR reduc* OR impair* OR disorder* OR fail* OR los* OR deficit* OR stop* OR progress* OR improve*)))

S26 TI "mental perform*" OR AB "mental perform*"

S27 TI ((memory N3 (complain* or declin* or function*))) OR AB ((memory N3 (complain* or declin* or function*)))

S28 S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27

S29 S6 AND S15 AND S28

Appendix 3: Chapter 3, completed PROBAST forms for included studies

First part of PROBAST form, applicable to all included studies

(Prediction model study Risk Of Bias Assessment Tool)

Published in Annals of Internal Medicine (freely available):

1. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies
2. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration

What does PROBAST assess?
PROBAST assesses both the *risk of bias* and *concerns regarding applicability* of a study that evaluates (develops, validates or updates) a multivariable diagnostic or prognostic prediction model. It is designed to assess primary studies included in a systematic review.

Bias occurs if systematic flaws or limitations in the design, conduct or analysis of a primary study distort the results. For the purpose of prediction modelling studies, we have defined *risk of bias* to occur when shortcomings in the study design, conduct or analysis lead to systematically distorted estimates of a model's predictive performance or to an inadequate model to address the research question. Model predictive performance is typically evaluated using calibration, discrimination and sometimes classification measures, and these are likely inaccurately estimated in studies with high risk of bias. *Applicability* refers to the extent to which the prediction model from the primary study matches your systematic review question, for example in terms of the participants, predictors or outcome of interest.

A primary study may include the development and/or validation or update of more than one prediction model. A PROBAST assessment should be completed for each distinct model that is developed, validated or updated (extended) for making individualised predictions. Where a publication assesses multiple prediction models, only complete a PROBAST assessment for those models that meet the inclusion criteria for your systematic review. Please note that subsequent use of the term "model" includes derivatives of models, such as simplified risk scores, nomograms, or recalibrations of models.

PROBAST is not designed for all multivariable diagnostic or prognostic studies. For example, studies using multivariable models to identify predictors associated with an outcome but not attempting to develop a model for making individualised predictions are not covered by PROBAST.

PROBAST includes four steps.

Step	Task	When to complete
1	Specify your systematic review question(s)	Once per systematic review
2	Classify the type of prediction model evaluation	Once for each model of interest in each publication being assessed, for each relevant outcome
3	Assess risk of bias and applicability	Once for each development and validation of each distinct prediction model in a publication
4	Overall judgment	Once for each development and validation of each distinct prediction model in a publication

If this is your first time using PROBAST, we strongly recommend reading the detailed explanation and elaboration (E&E, see link above) paper and to check the examples on www.probast.org

Step 1: Specify your systematic review question

State your systematic review question to facilitate the assessment of the applicability of the evaluated models to your question. *The following table should be completed once per systematic review.*

Criteria	Specify your systematic review question
Intended use of model:	Prognosis of cognitive outcome following stroke
***Participants** including selection criteria and setting:*	Adults with ischaemic stroke, haemorrhagic stroke, or TIA
***Predictors** (used in prediction modelling), including types of predictors (e.g. history, clinical examination, biochemical markers, imaging tests), time of measurement, specific measurement issues (e.g., any requirements/ prohibitions for specialized equipment):*	Demographics, medical history, lifestyle factors, clinical examination, biochemical markers, imaging data. All predictors should be collected or refer to a point in time preceding the occurrence of the cognitive outcome.
Outcome to be predicted:	Any cognitive outcome, including cognitive change, cognitive impairment, delirium, dementia, and cognitive recovery.

Second part of PROBAST form, completed individually for each included study

Chander et al., 2017; CHANGE score

Step 2: Classify the type of prediction model evaluation

Use the following table to classify the evaluation as model development, model validation or model update, or combination. Different signalling questions apply for different types of prediction model evaluation. If the evaluation does not fit one of these classifications then PROBAST should not be used.

Classify the evaluation based on its aim			
Type of prediction study	**PROBAST boxes to complete**	**Tick as appropriate**	**Definition for type of prediction model study**
Development only	Development	X	Prediction model development without external validation. These studies may include internal validation methods, such as bootstrapping and cross-validation techniques.
Development and validation	Development and validation	✓	Prediction model development combined with external validation in other participants in the same article.
Validation only	Validation	X	External validation of existing (previously developed) model in other participants.

This table should be completed once for each publication being assessed and for each relevant outcome in your review.

Publication reference	Chander, R. J. et al.; Development and validation of a risk score (CHANGE) for cognitive impairment after ischemic stroke; 2017
Models of interest	CHANGE score based on demographics and imaging variables
Outcome of interest	Post-stroke cognitive impairment

Step 3: Assess risk of bias and applicability

PROBAST is structured as four key domains. Each domain is judged for risk of bias (low, high or unclear) and includes signalling questions to help make judgements. Signalling questions are rated as yes (Y), probably yes (PY), probably no (PN), no (N) or no information (NI). All signalling questions are phrased so that “yes” indicates absence of bias. Any signalling question rated as “no” or “probably no” flags the potential for bias; you will need to use your

judgement to determine whether the domain should be rated as “high”, “low” or “unclear” risk of bias. The guidance document contains further instructions and examples on rating signalling questions and risk of bias for each domain. The first three domains are also rated for concerns regarding applicability (low/ high/ unclear) to your review question defined above.
Complete all domains separately for each evaluation of a distinct model. Shaded boxes indicate where signalling questions do not apply and should not be answered.

DOMAIN 1: Participants

A. Risk of Bias

Describe the sources of data and criteria for participant selection:

Development:
Retrospective cohort
Participants: diagnosis of ischaemic stroke; exclusion criteria: Discharge mRS >2; Subjects with pre-stroke cognitive impairment, neurological or psychiatric comorbidities, presented outside of the window period, or were unable to undergo cognitive assessments due to severe communication or visual disturbances as a result of the stroke, were excluded. Subjects with significant depression, screened via the 9-point Patient Health Questionnaire (PHQ-9) were also excluded.
Patients who were assessed by the clinical teams as being at risk for developing PSCI were scheduled for outpatient follow-up within 3–6 months after incident stroke.

Validation:
Retrospective cohort (STRIDE study); participants: inclusion criteria were: Chinese ethnicity, fluency in Cantonese, ability to participate in cognitive assessments, and provision of signed informed consent. Exclusion criteria for this study were: severe language impairment precluding cognitive assessment, presence of terminal illness, clinically significant psychiatric comorbidity, or known history of dementia before the index stroke. Severe language impairment was defined as a score of 3 points (i.e., mute) in the language score of the National Institute of Health Stroke Scale (NIHSS). All patients with stroke/TIA were invited to return for a neuropsychological assessment at 3 to 6 months after the index event.

		Dev	Val
1.1 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?		PY	PY
1.2 Were all inclusions and exclusions of participants appropriate?		N	Y
Risk of bias introduced by selection of participants	**RISK:** *(low/ high/ unclear)*	**H**	**L**

Rationale of bias rating:

Development:
It seems that all data for the study was collected at a single time-point, using admission records, which indicates that some predictors were not recorded for the purpose of research but routine clinical care - it is difficult to determine the quality of the data. However, medical records are overall considered a reliable source of information on general health-related variables. This is also not a factor that would necessarily overestimate the

performance of the model/lead to overoptimistic results. Moreover, routine clinical data is what would most likely be used if the model were applied in practice.
Exclusion criteria may have led to a non-representative patient sample, as stroke is likely to lead to disability (particularly as many patients may have had prior disability) and impairments impeding cognitive assessments, however, authors have stated that they wanted to focus on non-disabling strokes; excluding participants with depression may have lead to overestimation of prognostic accuracy, as it might be particularly difficult to distinguish between symptoms of depression and cognitive impairment, leading to an increase of false positives. High risk of bias due to enrolment being based on the clinical team's decision about who is at highest risk of PSCI, with no indication of how this decision was reached.

Validation:
As above, there are potential issues related to admission data not being recorded for research purposes, however this does not likely introduce a high risk of bias. This study was overall more inclusive, if exclusion criteria mainly applied where cognitive assessment would not be appropriate or possible, or where findings could be contributed to other significant medical conditions.

B. Applicability

Describe included participants, setting and dates:

Development:
Participants with non-disabling stroke who were found by the clinical team to be at risk of developing PSCI, recruitment from a tertiary centre in Singapore from Jan 2008 to Dec 2012

Validation:
Participant selection criteria as described above, recruitment from acute stroke unit in Hong Kong from Jan 2009 and Dec 2010

Concern that the included participants and setting do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	**L**

Rationale of applicability rating:

Development:
Sample may not be representative due to narrowing to patients assessed as likely to develop PSCI and selection criteria excluding disabled and depressed participants, with both disability and depression being relatively common in stroke populations; also potential cultural, geographical and healthcare differences, most obvious issue related to education – in western countries education <6 years would be an uncommon occurrence. However, the only criterium regarding participants was being an adult, diagnosed with stroke, so study population matches review question.

Validation:
Included participants match the review question, although potential cultural, geographical and healthcare differences, most obvious issue related to education – in western countries education <6 years would be an uncommon occurrence

DOMAIN 2: Predictors			
A. Risk of Bias			
List and describe predictors included in the final model, e.g. definition and timing of assessment: Predictors: Chronic lacunes, white matter hyperintensities, age, non-lacunar acute cortical infarcts, global cortical atrophy, education. Development: demographic, clinical and risk factor data were obtained from admission records, clinical MR images obtained at the time of stroke appraised by neurologist and neuroradiologist, and visually rated by blinded raters. Validation: demographic and clinical data collected during acute hospitalisation, MRI performed within first week of admission, appraised by 3 trained neurologists			
		Dev	Val
2.1 Were predictors defined and assessed in a similar way for all participants?		Y	Y
2.2 Were predictor assessments made without knowledge of outcome data?		PY	NI
2.3 Are all predictors available at the time the model is intended to be used?		Y	Y
Risk of bias introduced by predictors or their assessment	**RISK:** *(low/ high/ unclear)*	**L**	**L**
Rationale of bias rating: In development study, MR image raters were blinded and trained specialists, no information regarding blinding when assessing demographics, however these predictors do not involve subjective judgement. In validation study, MR images were appraised by trained specialists; blinding was not explicitly stated, however predictors were assessed prior to outcome.			
B. Applicability			
Concern that the definition, assessment or timing of predictors in the model do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	**L**
Rationale of applicability rating: There are potential issues with access to MRI, trained specialists using the assessment tools and subjectivity in rating MR images. However, everything matches review question and the setting for use was not specifically defined, e.g. these methods would still be highly applicable to research settings			

DOMAIN 3: Outcome

A. Risk of Bias

Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination:

Development:
"Cognitive status was assessed by clinical team via structured clinical interview and the Mini-Mental State Examination (MMSE). If further confirmation was required, the Singaporean version of the Montreal Cognitive Assessment (MoCA) was conducted. Subjects were classified as having PSCI if they had an MRI-confirmed infarct, met criteria for vascular cognitive impairment, and had MMSE ≤ 25 or MoCA ≤ 22. Remaining subjects were classified as having no cognitive impairment (NCI)."

Validation:
"All patients with stroke/TIA were invited to return for a neuropsychological assessment at 3 to 6 months after the index event. Trained psychologists administered the Clinical Dementia Rating (CDR), Cantonese Mini-Mental State Examination (MMSE), and Hong Kong version of the Montreal Cognitive Assessment to evaluate patients' cognitive functions while blinded to neuroimaging findings. The Chinese Geriatric Depression Scale was used to assess the extent of depressive symptoms. When grading using the CDR, care was taken particularly to grade only those impairments that were attributed to cognitive symptoms, not to motor/mood disturbances. Patients suspected of having dementia as defined by a CDR rating of 1 point or more at screening, and their caregivers, were invited for a detailed clinical assessment by neurologists specialized in dementia (V.C.T.M. and L.A.), who then confirmed a diagnosis of dementia according to the Diagnostic and Statistical Manual of Mental Disorders, fourth edition. During this clinical assessment, the neurologists also inquired about patients' cognitive function before the index event to exclude patients with dementia before the index event."

		Dev	Val
3.1 Was the outcome determined appropriately?		Y	Y
3.2 Was a pre-specified or standard outcome definition used?		PY	Y
3.3 Were predictors excluded from the outcome definition?		PN	Y
3.4 Was the outcome defined and determined in a similar way for all participants?		N	N?
3.5 Was the outcome determined without knowledge of predictor information?		PN	Y
3.6 Was the time interval between predictor assessment and outcome determination appropriate?		Y	Y
Risk of bias introduced by the outcome or its determination	**RISK:** *(low/ high/ unclear)*	**H**	**L**

Rationale of bias rating:

Development:
3.1 MMSE is considered to not have sufficient sensitivity to capture MCI in stroke populations, however it is nonetheless a standard, accepted, widely-used measure of cognition, 3.2 Not clear how criteria for VCI were applied (were definite, probable and possible all treated the same?), 3.3 Imaging variables (potentially predictors) are used to assess VCI, 3.4 some participants were additionally assessed with the MoCA without an objective indication of

who needed to be (seems to be based on subjective call), also there was a relatively big difference in assessment times (between 3 to 6 months) 3.5 given the diagnosing of VCI, it appears likely that assessors had some knowledge of imaging results.

Validation:
Seems that all participants were assessed using the same 3 standard measures, although not quite clear (possibly MoCA and MMSE were not always used in combination) and once again there was a relatively big difference in assessment times - 3 to 6 months (but does this go under 3.4 or 3.6?), CDR criteria are pre-specified and clear compared to what criteria might have been applied to diagnose VCI and also do not require knowledge of imaging findings, assessors were blinded to the latter.

B. Applicability

At what time point was the outcome determined:

Between 3 to 6 months

If a composite outcome was used, describe the relative frequency/distribution of each contributing outcome:

N/A

Concern that the outcome, its definition, timing or determination do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	**L**

Rationale of applicability rating:

The time interval was appropriate to capture PSCI and the review question did not impose and specific time restrictions regarding outcome assessment. The review focuses on any cognitive outcome assessed using a recognised measure, and therefore the chosen outcome matches the review question.

DOMAIN 4: Analysis

Risk of Bias

Describe numbers of participants, number of candidate predictors, outcome events and events per candidate predictor:

Development: 209 participants, 26 candidate predictors (intracranial stenosis was not mentioned in the table, but was in the text), 78 events, 3.0 events per predictor

Validation: 693 participants, 352 with PSCI

Describe how the model was developed (for example in regards to modelling technique (e.g. survival or logistic modelling), predictor selection, and risk group definition):

Logistic regression
Variables were deemed eligible for inclusion in the initial stage of model building if they were: 1) statistically significant at the univariate level after operationalization, 2) found in the literature to be relevant, and 3) were deemed by the study team that the variables were common enough to be available to existing stroke workflows
“Statistically significant continuous variables were operationalized into categorical variables based on clinically relevant cutoffs and retested for statistical significance.”

Describe whether and how the model was validated, either internally (e.g. bootstrapping, cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different setting, different type of participants):

Apparent and temporal with stability assessed using 10-fold cross-validation; external

Describe the performance measures of the model, e.g. (re)calibration, discrimination, (re)classification, net benefit, and whether they were adjusted for optimism:

Discrimination, classification abilities, calibration, cross-validation to assess stability of AUROC estimates

Describe any participants who were excluded from the analysis:

Development: 34 excluded – 6 presented outside of 3-6 months, 28 had incomplete investigative data

Validation: 314 subjects were excluded - 71 with hemorrhagic stroke, 140 with TIA, 86 with non-stroke or unknown etiologies, and 17 with history of intracranial hemorrhage); also participants with incomplete data were excluded from the STRIDE study itself

Describe missing data on predictors and outcomes as well as methods used for missing data:

Information on missing data was not presented, participants with missing data were excluded from the analysis

		Dev	Val
4.1	Were there a reasonable number of participants with the outcome?	N	Y
4.2	Were continuous and categorical predictors handled appropriately?	NI	PY
4.3	Were all enrolled participants included in the analysis?	N	N
4.4	Were participants with missing data handled appropriately?	N	N
4.5	Was selection of predictors based on univariable analysis avoided?	N	
4.6	Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?	PN	N
4.7	Were relevant model performance measures evaluated appropriately?	Y	Y
4.8	Were model overfitting and optimism in model performance accounted for?	Y	
4.9	Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?	N	
Risk of bias introduced by the analysis	**RISK:** *(low/ high/ unclear)*	**H**	**H**

Rationale of bias rating:

Development:
Too few participants with event relative to candidate predictors, exclusion of participants with missing data, selection of predictors based on univariable analysis, global cortical atrophy and white matter hyperintensities were

included in the risk score despite not being significant in multivariable regression – inconsistency in approach Validation: Issues due to missing data and exclusion of participants lost to follow-up/dead.

Step 4: Overall assessment

Use the following tables to reach overall judgements about risk of bias and concerns regarding applicability of the prediction model evaluation (development and/or validation) across all assessed domains.
Complete for each evaluation of a distinct model.

Reaching an overall judgement about risk of bias of the prediction model evaluation	
Low risk of bias	If all domains were rated low risk of bias. If a prediction model was developed without any external validation, and it was rated as low risk of bias for all domains, consider downgrading to **high risk of bias**. Such a model can only be considered as low risk of bias, if the development was based on a very large data set and included some form of internal validation.
High risk of bias	If at least one domain is judged to be at **high risk of bias**.
Unclear risk of bias	If an unclear risk of bias was noted in at least one domain and it was low risk for all other domains.

Reaching an overall judgement about applicability of the prediction model evaluation	
Low concerns regarding applicability	If low concerns regarding applicability for all domains, the prediction model evaluation is judged to have **low concerns regarding applicability**.
High concerns regarding applicability	If high concerns regarding applicability for at least one domain, the prediction model evaluation is judged to have **high concerns regarding applicability**.
Unclear concerns regarding applicability	If unclear concerns (but no "high concern") regarding applicability for at least one domain, the prediction model evaluation is judged to have **unclear concerns regarding applicability** overall.

Overall judgement about risk of bias and applicability of the prediction model evaluation		
Overall judgement of risk of bias	**RISK:** *(low/ high/ unclear)*	**H**
Summary of sources of potential bias: Assessment of outcome associated with knowledge of predictors and differing across subject; issues with model development, including type of analysis, selection of predictors and handling missing data		
Overall judgement of applicability	**CONCERN:** *(low/ high/ unclear)*	**L**
Summary of applicability concerns: selection criteria that exclude participants with common sequalae of stroke, including depression and any disability resulting in dependence, however the study population nonetheless matches the review question		

Ding et al., 2019

Step 2: Classify the type of prediction model evaluation

Use the following table to classify the evaluation as model development, model validation or model update, or combination. Different signalling questions apply for different types of prediction model evaluation. If the evaluation does not fit one of these classifications then PROBAST should not be used.

Classify the evaluation based on its aim			
Type of prediction study	**PROBAST boxes to complete**	**Tick as appropriate**	**Definition for type of prediction model study**
Development only	Development	✓	Prediction model development without external validation. These studies may include internal validation methods, such as bootstrapping and cross-validation techniques.
Development and validation	Development and validation	X	Prediction model development combined with external validation in other participants in the same article.
Validation only	Validation	X	External validation of existing (previously developed) model in other participants.

This table should be completed once for each publication being assessed and for each relevant outcome in your review.

Publication reference	Ding, M. Y. et al. Predictors of Cognitive Impairment after Stroke: A Prospective Stroke Cohort Study. 2019
Models of interest	Mathematical formula using demographic, clinical and imaging variables
Outcome of interest	Post-stroke cognitive impairment diagnosed based on MMSE, MoCA, CDR and DSM-4 criteria

Step 3: Assess risk of bias and applicability

PROBAST is structured as four key domains. Each domain is judged for risk of bias (low, high or unclear) and includes signalling questions to help make judgements. Signalling questions are rated as yes (Y), probably yes (PY), probably no (PN), no (N) or no information (NI). All signalling questions are phrased so that "yes" indicates absence of bias. Any signalling question rated as "no" or "probably no" flags the potential for bias; you will need to use your judgement to determine whether the domain should be rated as "high", "low" or "unclear" risk of bias. The guidance document contains further instructions and examples on rating signalling questions and risk of bias for each domain. The first three domains are also rated for concerns regarding applicability (low/ high/ unclear) to your review question defined above.
Complete all domains separately for each evaluation of a distinct model. Shaded boxes indicate where signalling questions do not apply and should not be answered.

DOMAIN 1: Participants			
A. Risk of Bias			
Describe the sources of data and criteria for participant selection: Prospective cohort; Inclusion criteria: diagnosis of stroke, stroke onset within last 7 days, adult, willingness to undergo assessments and data collection procedures; Exclusion criteria: pregnancy, severe vital organ failure, major mental illness, history of dementia, participation in other trials			
		Dev	Val
1.3 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?		Y	
1.4 Were all inclusions and exclusions of participants appropriate?		Y	
Risk of bias introduced by selection of participants	**RISK:** *(low/ high/ unclear)*	**L**	
Rationale of bias rating: Exclusion criteria seem appropriate as future cognitive outcomes are less of a priority where patient's condition may be terminal and are less likely to be assessed in a real-world healthcare setting, inclusion of patients with dementia may have led to overoptimistic model estimations, outcomes of patients taking part in other trials may be affected by interventions.			
B. Applicability			
Describe included participants, setting and dates: Tertiary care centre in China, patients diagnosed with acute ischaemic stroke; data collection: June 2017 to May 2018			

Concern that the included participants and setting do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating: Although applicability may be affected by cultural, geographical and healthcare settings, the study sample seems representative of stroke population of interest			

DOMAIN 2: Predictors			
A. Risk of Bias			
List and describe predictors included in the final model, e.g. definition and timing of assessment: Age, years of education, periventricular hyperintensity grading (PVH), diabetes mellitus (DM), number of acute nonlacunar infarcts. DM was defined by fasting plasma glucose, 2h postprandial glucose or treatment with insulin or oral hypoglycemic medication. Acute nonlacunar infarct defined as lesion greater than 20mm. PVH assessed based on Fazekas scale.			
		Dev	Val
2.4 Were predictors defined and assessed in a similar way for all participants?		N	
2.5 Were predictor assessments made without knowledge of outcome data?		NI	
2.6 Are all predictors available at the time the model is intended to be used?		Y	
Risk of bias introduced by predictors or their assessment	**RISK:** *(low/ high/ unclear)*	**L**	
Rationale of bias rating: Although different measures were used to determine DM, all seem valid, particularly that lab findings might not indicate disease in patients taking medication, while some patients might be undiagnosed or not take medication for other reasons; as the study is prospective it can be inferred the predictors were assessed without knowledge of the outcome; it is not stated whether MR image raters had similar levels of expertise and assessment of these predictors allows for subjectivity, however the included predictors were well-defined and based on recognised, standard tool.			
B. Applicability			
Concern that the definition, assessment or timing of predictors in the model do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating: Definition, assessment and timing match our broad review question			

DOMAIN 3: Outcome

A. Risk of Bias

Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination:

All patients were asked to complete the MMSE, MoCA, a neuropsychological assessment battery and were assessed based on CDR and DSM-4 criteria. PSCI identified if both MMSE and MoCA scores were lower than cut-off or a neuropsychological battery with more than one affected domain if the MMSE and MoCA were discordant and the CDR rating was greater than 0 points. Post-stroke dementia was diagnosed based on DSM-4 criteria. Follow-up was between 6 and 12 months

		Dev	Val
3.7 Was the outcome determined appropriately?		PY	
3.8 Was a pre-specified or standard outcome definition used?		Y	
3.9 Were predictors excluded from the outcome definition?		Y	
3.10 Was the outcome defined and determined in a similar way for all participants?		N?	
3.11 Was the outcome determined without knowledge of predictor information?		NI	
3.12 Was the time interval between predictor assessment and outcome determination appropriate?		Y	
Risk of bias introduced by the outcome or its determination	**RISK:** *(low/ high/ unclear)*	**UN**	

Rationale of bias rating:

This is difficult to assess without knowing whether assessors had knowledge of predictor information, as assessment based on CDR and DSM-4 allows for subjective judgement, while examiner may influence test results with their behaviour, potentially large difference in follow-up time between participants (6 to 12 months); the level of assessor expertise is also unclear.

B. Applicability

At what time point was the outcome determined:

Between 6 to 12 months

If a composite outcome was used, describe the relative frequency/distribution of each contributing outcome:

N/A

Concern that the outcome, its definition, timing or determination do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	

Rationale of applicability rating:

Although it is unlikely that cognition will be routinely assessed using a method as complex as described in this study, it seems credible (likely to correctly identify cases of PSCI); all aspects of the outcome match the review question.

DOMAIN 4: Analysis			
Risk of Bias			
Describe numbers of participants, number of candidate predictors, outcome events and events per candidate predictor: 145, 77 outcome events, number of candidate predictors is unclear but seems like 52, and at least 45, so EPV at most 1.7			
Describe how the model was developed (for example in regards to modelling technique (e.g. survival or logistic modelling), predictor selection, and risk group definition): Logistic regression, predictors selected based on results of univariable analysis			
Describe whether and how the model was validated, either internally (e.g. bootstrapping, cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different setting, different type of participants): Paper mentions bootstrap in estimating predictor ORs, but doesn't seem to have been applied to AUROC estimate, also unclear on number of replications, seems like only apparent validation was done.			
Describe the performance measures of the model, e.g. (re)calibration, discrimination, (re)classification, net benefit, and whether they were adjusted for optimism: Discrimination with AUROC, calibration, but only based on Hosmer-Lemeshow test			
Describe any participants who were excluded from the analysis: 34 subjects did not participate in follow-up visit due to rehabilitation in admission, not in Shanghai, refusal.			
Describe missing data on predictors and outcomes as well as methods used for missing data: No indication of any missing data			
		Dev	Val
4.10 Were there a reasonable number of participants with the outcome?		N	
4.11 Were continuous and categorical predictors handled appropriately?		Y	
4.12 Were all enrolled participants included in the analysis?		N	
4.13 Were participants with missing data handled appropriately?		NI	
4.14 Was selection of predictors based on univariable analysis avoided?		PN	
4.15 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?		N	
4.16 Were relevant model performance measures evaluated appropriately?		N	
4.17 Were model overfitting and optimism in model performance accounted for?		N	
4.18 Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?		NI	
Risk of bias introduced by the analysis	**RISK:** *(low/ high/ unclear)*	**H**	

Rationale of bias rating:

Too few participants relative to number of candidate predictors, participants lost to follow-up were excluded, method of handling missing data not explicitly stated, predictors assessed based on univariable associations, no appropriate measure of calibration, only apparent validation with AUROC no corrected for optimism, also authors didn't report on results of logistic regression models or even what exact variables were included in them, therefore it is not possible to assess 4.9.

Step 4: Overall assessment

Use the following tables to reach overall judgements about risk of bias and concerns regarding applicability of the prediction model evaluation (development and/or validation) across all assessed domains.
Complete for each evaluation of a distinct model.

Reaching an overall judgement about risk of bias of the prediction model evaluation	
Low risk of bias	If all domains were rated low risk of bias. If a prediction model was developed without any external validation, and it was rated as low risk of bias for all domains, consider downgrading to **high risk of bias**. Such a model can only be considered as low risk of bias, if the development was based on a very large data set and included some form of internal validation.
High risk of bias	If at least one domain is judged to be at **high risk of bias**.
Unclear risk of bias	If an unclear risk of bias was noted in at least one domain and it was low risk for all other domains.

Reaching an overall judgement about applicability of the prediction model evaluation	
Low concerns regarding applicability	If low concerns regarding applicability for all domains, the prediction model evaluation is judged to have **low concerns regarding applicability**.
High concerns regarding applicability	If high concerns regarding applicability for at least one domain, the prediction model evaluation is judged to have **high concerns regarding applicability**.
Unclear concerns regarding applicability	If unclear concerns (but no "high concern") regarding applicability for at least one domain, the prediction model evaluation is judged to have **unclear concerns regarding applicability** overall.

Overall judgement about risk of bias and applicability of the prediction model evaluation		
Overall judgement of risk of bias	**RISK:** *(low/ high/ unclear)*	**H**
Summary of sources of potential bias: High risk of bias in terms of analysis due to exclusion of participants lost to follow-up, type of analysis, selection of predictors, and estimation of model performance; also unclear whether outcome was assessed with blinding to predictors		
Overall judgement of applicability	**CONCERN:** *(low/ high/ unclear)*	**L**
Summary of applicability concerns: Study sample seems likely to be representative of stroke population in which cognitive outcome is recognised as relevant/important (not terminally ill or with level of functional disability requiring full-time care), all aspects match review question.		

Gong et al., 2019

Step 2: Classify the type of prediction model evaluation

Use the following table to classify the evaluation as model development, model validation or model update, or combination. Different signalling questions apply for different types of prediction model evaluation. If the evaluation does not fit one of these classifications then PROBAST should not be used.

Classify the evaluation based on its aim			
Type of prediction study	**PROBAST boxes to complete**	**Tick as appropriate**	**Definition for type of prediction model study**
Development only	Development	✓	Prediction model development without external validation. These studies may include internal validation methods, such as bootstrapping and cross-validation techniques.
Development and validation	Development and validation	X	Prediction model development combined with external validation in other participants in the same article.
Validation only	Validation	X	External validation of existing (previously developed) model in other participants.

This table should be completed once for each publication being assessed and for each relevant outcome in your review.

Publication reference	Gong, K. A nomogram to predict cognitive impairment after supratentorial spontaneous intracranial hematoma in adult patients: A retrospective cohort study. 2019
Models of interest	Nomogram based on Glasgow Coma Scale, bleeding volume and presence of intraventricular haemorrhage
Outcome of interest	Cognitive impairment after supratentorial haemorrhage

Step 3: Assess risk of bias and applicability

PROBAST is structured as four key domains. Each domain is judged for risk of bias (low, high or unclear) and includes signalling questions to help make judgements. Signalling questions are rated as yes (Y), probably yes (PY), probably no (PN), no (N) or no information (NI). All signalling questions are phrased so that "yes" indicates absence of bias. Any signalling question rated as "no" or "probably no" flags the potential for bias; you will need to use your judgement to determine whether the domain should be rated as "high", "low" or "unclear" risk of bias. The guidance document contains further instructions and examples on rating signalling questions and risk of bias for each domain. The first three domains are also rated for concerns regarding applicability (low/ high/ unclear) to your review question defined above.
Complete all domains separately for each evaluation of a distinct model. Shaded boxes indicate where signalling questions do not apply and should not be answered.

DOMAIN 1: Participants			
A. Risk of Bias			
Describe the sources of data and criteria for participant selection: Retrospective cohort 170 consecutive patients with intracranial hematoma were enrolled in this study. All participants were admitted in rehabilitation department and had received initial treatment at the onset of stroke in 1 hospital in southeast area of China. Eleven patients were excluded due to dementia past history. Nineteen patients with stroke history before and 23 patients with subtentorial hematoma were also excluded. We divided the remaining 127 patients into 2 datasets according their admission time, 92 patients who were admitted before 2018 were enrolled in development dataset.			
		Dev	Val
1.5 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?		Y	
1.6 Were all inclusions and exclusions of participants appropriate?		Y	
Risk of bias introduced by selection of participants	**RISK:** *(low/ high/ unclear)*	**L**	
Rationale of bias rating: Participants with previous stroke are often excluded due to potential confounding from pre-existing disability and cognitive impairment, inclusion of patients with dementia may lead to overoptimistic model estimations, the aim of the study was to focus only on patients with haemorrhage.			

B. Applicability			
Describe included participants, setting and dates: Patients with intracranial hematoma, 1 hospital in southeast area of China; January 1, 2016 to October 31, 2018			
Concern that the included participants and setting do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating: Study limited to one particular stroke type, however this was the specified aim and still fits with the review question, applicability may be affected by cultural, geographical and healthcare settings			

DOMAIN 2: Predictors			
A. Risk of Bias			
List and describe predictors included in the final model, e.g. definition and timing of assessment: Bleeding volume and intraventricular haemorrhage evaluated independently by 2 neurosurgeons. GCS assessed at stroke onset by clinicians in charge of patient care.			
		Dev	Val
2.7 Were predictors defined and assessed in a similar way for all participants?		Y	
2.8 Were predictor assessments made without knowledge of outcome data?		NI	
2.9 Are all predictors available at the time the model is intended to be used?		Y	
Risk of bias introduced by predictors or their assessment	**RISK:** *(low/ high/ unclear)*	**L**	
Rationale of bias rating: 2.2 is unclear, assessment of CT images allows for some subjectivity, however predictors seem well-defined and 2 independent raters were involved.			
B. Applicability			
Concern that the definition, assessment or timing of predictors in the model do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating: All aspects related to predictors match the review question.			

DOMAIN 3: Outcome

A. Risk of Bias

Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination:

Mini-mental State Examination (MMSE) score was evaluated after 3 months from the onset of the disease by 1 qualified rehabilitation physician. MMSE ≤ 24 was considered as cognitive impairment.

		Dev	Val
3.13 Was the outcome determined appropriately?		Y	
3.14 Was a pre-specified or standard outcome definition used?		Y	
3.15 Were predictors excluded from the outcome definition?		Y	
3.16 Was the outcome defined and determined in a similar way for all participants?		Y	
3.17 Was the outcome determined without knowledge of predictor information?		NI	
3.18 Was the time interval between predictor assessment and outcome determination appropriate?		Y	
Risk of bias introduced by the outcome or its determination	**RISK:** *(low/ high/ unclear)*	**L**	

Rationale of bias rating:

MMSE is not as sensitive as e.g. MoCA that is recommended for stroke populations, yet it is nonetheless a recognised, validated and widely-used measure of cognitive function, 3.5 is unclear and test scores may be to a certain extent influenced by examiner behaviour, yet given the binary nature of the outcome, this would likely only affect participants with actual scores near the cut off, also the examiner seems qualified to carry out the assessment appropriately

B. Applicability

At what time point was the outcome determined:

3 months after stroke onset

If a composite outcome was used, describe the relative frequency/distribution of each contributing outcome:

N/A

Concern that the outcome, its definition, timing or determination do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	

Rationale of applicability rating:

All aspects of the outcome match the review question.

DOMAIN 4: Analysis

Risk of Bias

Describe numbers of participants, number of candidate predictors, outcome events and events per candidate predictor:

92 participants, 13 candidate predictors, 69 outcome events in development and validation datasets (unknown for development set alone), even if all 69 were in development set, EPV = 5.3

Describe how the model was developed (for example in regards to modelling technique (e.g. survival or logistic modelling), predictor selection, and risk group definition):

We used logistic regression to build models on the development dataset. Cognitive impairment was chosen to be the response variable. Age, gender, hemorrhage sites, bleeding volume, GCS, IVH, DM, hypertension, hyperlipidemia were potential risk factors. We classified GCS into 3 categories: GCS ≥ 12, 9 ≤ GCS < 12 and GCS ≤ 8. Bleeding volume was also divided into ordered categorical factors as volume < 10 ml, 10 ml ≤ volume < 20 ml, 20 ml ≤ volume < 30 ml, 30 ml ≤ volume < 40 ml, 40 ml ≤ volume < 50 ml, 50 ml ≤ volume < 60 ml and volume ≥ 60 ml. Risk factors which showed statistically significant in univariate logistic analysis would be enrolled in multiple logistic regression. Akaike Information Criterion (AIC) was used to determine which factors should be enrolled in the final model.

Describe whether and how the model was validated, either internally (e.g. bootstrapping, cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different setting, different type of participants):

We divided the remaining 127 patients into 2 datasets according their admission time, 92 patients who were admitted before 2018 were enrolled in development dataset. The other 35 patients were in validation dataset.

Describe the performance measures of the model, e.g. (re)calibration, discrimination, (re)classification, net benefit, and whether they were adjusted for optimism:

A nomogram was drawn according to the final logistic regression model. Receiver operating characteristic (ROC) curve and calibration curve on development and validation datasets were drawn separately.

Describe any participants who were excluded from the analysis:

Participants excluded as described in Domain A, apparently no loss to follow up

Describe missing data on predictors and outcomes as well as methods used for missing data:

No mention of missing data

	Dev	Val
4.19 Were there a reasonable number of participants with the outcome?	N	
4.20 Were continuous and categorical predictors handled appropriately?	PN	
4.21 Were all enrolled participants included in the analysis?	Y	
4.22 Were participants with missing data handled appropriately?	NI	
4.23 Was selection of predictors based on univariable analysis avoided?	N	

4.24 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?		NI	
4.25 Were relevant model performance measures evaluated appropriately?		Y	
4.26 Were model overfitting and optimism in model performance accounted for?		N	
4.27 Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?		PY	
Risk of bias introduced by the analysis	**RISK:** *(low/ high/ unclear)*	**H**	
Rationale of bias rating: Too few participants for model development and validation, GCS categorised differently than recommended (score of 12 should be treated as more severe) and uncertain how many patients with what variability In volume were grouped in bleeding volume >60, seletion of predictors based on univariable associations, inappropriate method of internal validation.			

Step 4: Overall assessment

Use the following tables to reach overall judgements about risk of bias and concerns regarding applicability of the prediction model evaluation (development and/or validation) across all assessed domains.
Complete for each evaluation of a distinct model.

Reaching an overall judgement about risk of bias of the prediction model evaluation	
Low risk of bias	If all domains were rated low risk of bias. If a prediction model was developed without any external validation, and it was rated as low risk of bias for all domains, consider downgrading to **high risk of bias**. Such a model can only be considered as low risk of bias, if the development was based on a very large data set and included some form of internal validation.
High risk of bias	If at least one domain is judged to be at **high risk of bias**.
Unclear risk of bias	If an unclear risk of bias was noted in at least one domain and it was low risk for all other domains.

Reaching an overall judgement about applicability of the prediction model evaluation	
Low concerns regarding applicability	If low concerns regarding applicability for all domains, the prediction model evaluation is judged to have **low concerns regarding applicability**.
High concerns regarding applicability	If high concerns regarding applicability for at least one domain, the prediction model evaluation is judged to have **high concerns regarding applicability**.
Unclear concerns regarding applicability	If unclear concerns (but no "high concern") regarding applicability for at least one domain, the prediction model evaluation is judged to have **unclear concerns regarding applicability** overall.

Overall judgement about risk of bias and applicability of the prediction model evaluation		
Overall judgement of risk of bias	**RISK:** *(low/ high/ unclear)*	**H**
Summary of sources of potential bias: Source of bias stems from model development and validation issues, including insufficient sample size, predictor selection and use of split-sample validation.		
Overall judgement of applicability	**CONCERN:** *(low/ high/ unclear)*	**L**
Summary of applicability concerns: All domains match review question, few exclusion criteria so sample likely to be representative.		

Kandiah et al., 2016; SIGNAL$_2$

Step 2: Classify the type of prediction model evaluation

Use the following table to classify the evaluation as model development, model validation or model update, or combination. Different signalling questions apply for different types of prediction model evaluation. If the evaluation does not fit one of these classifications then PROBAST should not be used.

Classify the evaluation based on its aim			
Type of prediction study	**PROBAST boxes to complete**	**Tick as appropriate**	**Definition for type of prediction model study**
Development only	Development	✓	Prediction model development without external validation. These studies may include internal validation methods, such as bootstrapping and cross-validation techniques.
Development and validation	Development and validation	X	Prediction model development combined with external validation in other participants in the same article.
Validation only	Validation	X	External validation of existing (previously developed) model in other participants.

This table should be completed once for each publication being assessed and for each relevant outcome in your review.

Publication reference	Kandiah, N. et al. Cognitive Impairment after Mild Stroke: Development and Validation of the SIGNAL2 Risk Score. 2015
Models of interest	SIGNAL2 risk score based on demographics and imaging variables.
Outcome of interest	Post-stroke cognitive impairment

Step 3: Assess risk of bias and applicability

PROBAST is structured as four key domains. Each domain is judged for risk of bias (low, high or unclear) and includes signalling questions to help make judgements. Signalling questions are rated as yes (Y), probably yes (PY), probably no (PN), no (N) or no information (NI). All signalling questions are phrased so that "yes" indicates absence of bias. Any signalling question rated as "no" or "probably no" flags the potential for bias; you will need to use your judgement to determine whether the domain should be rated as "high", "low" or "unclear" risk of bias. The guidance document contains further instructions and examples on rating signalling questions and risk of bias for each domain. The first three domains are also rated for concerns regarding applicability (low/ high/ unclear) to your review question defined above.
Complete all domains separately for each evaluation of a distinct model. Shaded boxes indicate where signalling questions do not apply and should not be answered.

DOMAIN 1: Participants			
A. Risk of Bias			
Describe the sources of data and criteria for participant selection: Retrospective cohort; "From this database, we included patients with mild ischemic stroke, defined as having a modified Rankin scale of ≤2 with an MR-confirmed acute infarct. Exclusion criteria included presence of pre-stroke cognitive impairment, absence of MRI performed during acute stroke work-up, presentation to outpatient clinic outside 3–6 months, neurological or psychiatric comorbidities, unable to undergo cognitive assessments due to severe communication or visual disturbances as a result of the stroke. Subjects with significant depression, screened via the 9-point Patient Health Questionnaire (PHQ-9) were also excluded." "Patients who were assessed by the clinical teams as being at risk for developing PSCI were scheduled for outpatient follow-up within 3–6 months after incident stroke."			
		Dev	Val
1.7 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?		Y	
1.8 Were all inclusions and exclusions of participants appropriate?		N	
Risk of bias introduced by selection of participants	**RISK:** *(low/ high/ unclear)*	**H**	

Rationale of bias rating:

It seems that all data for the study was collected at a single time-point, using admission records, which indicates that some predictors were not recorded for the purpose of research but routine clinical care - it is difficult to determine the quality of the data. However, medical records are overall considered a reliable source of information on general health-related variables. This is also not a factor that would necessarily overestimate the performance of the model/lead to overoptimistic results. Moreover, routine clinical data is what would most likely be used if the model were applied in practice.

Exclusion criteria may have led to a non-representative patient sample, as stroke is likely to lead to disability (particularly as many patients may have had prior disability) and impairments impeding cognitive assessments, however, authors have stated that they wanted to focus on non-disabling strokes; excluding participants with depression may have led to overestimation of prognostic accuracy, as it might be particularly difficult to distinguish between symptoms of depression and cognitive impairment, leading to an increase of false positives. High risk of bias due to enrolment being based on the clinical team's decision about who is at highest risk of PSCI, with no indication of how this decision was reached.

B. Applicability

Describe included participants, setting and dates:

Participants with non-disabling stroke who were found by the clinical team to be at risk of developing PSCI, recruitment from a tertiary centre in Singapore from Jan 2008 to Dec 2012

Concern that the included participants and setting do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	

Rationale of applicability rating:

Sample may not be representative due to narrowing to patients assessed as likely to develop PSCI and selection criteria excluding those disabled and depressed, with both disability and depression being relatively common in stroke populations (issues with model generalisability); also potential cultural, geographical and healthcare differences, most obvious issue related to education – in western countries education <6 years would be an unlikely occurrence. However, the only criterium regarding participants was being an adult, diagnosed with stroke, so study population matches review question.

DOMAIN 2: Predictors			
A. Risk of Bias			
List and describe predictors included in the final model, e.g. definition and timing of assessment: Age, education <6 years, global cortical atrophy stages, fazekas stages, nonlacunar cortical infarct stages, chronic lacunes >=2, intracranial stenosis. "Data on demographic factors and vascular risk factors were collected from admission records. MR images were reviewed independently by a neurologist and a neuroradiologist. Acute infarcts were quantified based on the number and location of lacunar and non-lacunar infarcts. T2 sequences were used to quantify white matter hyperintensity (WMH) using the Fazekas scale. Chronic lacunes were quantified by number and location. Gradient-echo sequences were used to rate microhemorrhages using the Microbleeds Anatomical Rating Scale. MRA images were rated for presence and severity of intracranial large vessel stenosis. T1 sequences were also assessed for global cortical atrophy (GCA). Any differences in ratings between raters were resolved by consensus."			
		Dev	Val
2.10 Were predictors defined and assessed in a similar way for all participants?		Y	
2.11 Were predictor assessments made without knowledge of outcome data?		NI	
2.12 Are all predictors available at the time the model is intended to be used?		Y	
Risk of bias introduced by predictors or their assessment	**RISK:** *(low/ high/ unclear)*	**L**	
Rationale of bias rating: It is unclear from this paper whether raters had knowledge of outcome data, however subsequent publication suggests that they were blinded. Also having two independent, well-qualified raters reduces risk of bias for assessing imaging variables, particularly as recognised and validated tools were applied.			
B. Applicability			
Concern that the definition, assessment or timing of predictors in the model do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating: There are potential issues with access to MRI, trained specialists using the assessment tools and subjectivity in rating MR images. However, everything matches review question and the setting for use was not specifically defined, e.g. these methods would still be highly applicable to research settings.			

DOMAIN 3: Outcome

A. Risk of Bias

Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination:

"During outpatient care, patients were assessed for progression of risk factors and for cognitive status using the Mini-Mental State Exam (MMSE). Where clinically indicated or further confirmation of cognitive status was required, the Montreal Cognitive Assessment (MoCA) was carried out as well.
Subjects in the development dataset were classified as PSCI if there were any cognitive symptoms during the clinical visit, and scored either MMSE ≤25 or MoCA ≤22, following validated local cutoffs. PSCI patients did not meet criteria for dementia. Patients who did not meet criteria for PSCI or dementia were categorized as "No Cognitive Impairment (NCI)"."

		Dev	Val
3.19 Was the outcome determined appropriately?		PY	
3.20 Was a pre-specified or standard outcome definition used?		PY	
3.21 Were predictors excluded from the outcome definition?		NI	
3.22 Was the outcome defined and determined in a similar way for all participants?		N	
3.23 Was the outcome determined without knowledge of predictor information?		NI	
3.24 Was the time interval between predictor assessment and outcome determination appropriate?		Y	
Risk of bias introduced by the outcome or its determination	**RISK:** *(low/ high/ unclear)*	**H**	

Rationale of bias rating:

The statement "if there were any cognitive symptoms during the clinical visit" does not allow to determine how these symptoms were assessed and what was considered sufficient to provide evidence of PSCI, there is no information regarding whether the outcome was determined without knowledge of predictors, however information from the subsequent paper suggests that predictors could have even actually considered in the process; some participants additionally completed the MoCA (a more sensitive measure) with selection criteria remaining unclear. Also, there was relatively big variability in the interval from index stroke to follow-up.

B. Applicability

At what time point was the outcome determined:

3 to 6 months post-stroke

If a composite outcome was used, describe the relative frequency/distribution of each contributing outcome:

N/A

Concern that the outcome, its definition, timing or determination do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	

Rationale of applicability rating:
Despite the relatively big variability in assessment timing, all aspects of the outcome match the review question.

DOMAIN 4: Analysis		
Risk of Bias		
Describe numbers of participants, number of candidate predictors, outcome events and events per candidate predictor: 209 participants, 43 candidate predictors, 78 outcome events, EPV = 1.8		
Describe how the model was developed (for example in regards to modelling technique (e.g. survival or logistic modelling), predictor selection, and risk group definition): "Potential variables for the predictive model were identified by comparing PSCI and NCI patients and tested for significance using independent sample *t*test or Wilcoxon-Mann-Whitney test for continuous data, and *x*2 test or Fisher's Exact test for categorical data. Statistically significant continuous variables were converted into categorical variables based on clinical cutoffs and retested using *x*2 test or Fisher's Exact test. All variables significant at univariate level were put into multivariate logistic regression models as predictor variables, with PSCI status as the outcome variable. A point system was developed based on the beta coefficients from the final model."		
Describe whether and how the model was validated, either internally (e.g. bootstrapping, cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different setting, different type of participants): Apparent and temporal with stability assessed using 10-fold cross-validation		
Describe the performance measures of the model, e.g. (re)calibration, discrimination, (re)classification, net benefit, and whether they were adjusted for optimism: Discrimination, calibration, and classification measures		
Describe any participants who were excluded from the analysis: "We excluded 34 patients (six presented outside of 3 to 6 months and 28 had incomplete clinical or investigative data)."		
Describe missing data on predictors and outcomes as well as methods used for missing data: Participants with missing data were excluded from the analysis.		
	Dev	Val
4.28 Were there a reasonable number of participants with the outcome?	N	
4.29 Were continuous and categorical predictors handled appropriately?	NI	
4.30 Were all enrolled participants included in the analysis?	N	
4.31 Were participants with missing data handled appropriately?	N	
4.32 Was selection of predictors based on univariable analysis avoided?	N	
4.33 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?	PN	
4.34 Were relevant model performance measures evaluated appropriately?	Y	
4.35 Were model overfitting and optimism in model performance accounted for?	Y	

4.36 Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?		N	
Risk of bias introduced by the analysis	**RISK:** *(low/ high/ unclear)*	**H**	
Rationale of bias rating: Too few participants relative to number of candidate predictors, continuous predictors apparently categorised based on clinical cut-offs, but no references provided, also unclear whether cut-offs were specified prior to analysis, predictors selected based on univariable analysis, participants with missing data excluded from analysis, based on how scores were assigned to age, education <6 should have been assigned 4 points.			

Step 4: Overall assessment

Use the following tables to reach overall judgements about risk of bias and concerns regarding applicability of the prediction model evaluation (development and/or validation) across all assessed domains.
Complete for each evaluation of a distinct model.

Reaching an overall judgement about risk of bias of the prediction model evaluation	
Low risk of bias	If all domains were rated low risk of bias. If a prediction model was developed without any external validation, and it was rated as low risk of bias for all domains, consider downgrading to **high risk of bias**. Such a model can only be considered as low risk of bias, if the development was based on a very large data set and included some form of internal validation.
High risk of bias	If at least one domain is judged to be at **high risk of bias**.
Unclear risk of bias	If an unclear risk of bias was noted in at least one domain and it was low risk for all other domains.

Reaching an overall judgement about applicability of the prediction model evaluation	
Low concerns regarding applicability	If low concerns regarding applicability for all domains, the prediction model evaluation is judged to have **low concerns regarding applicability**.
High concerns regarding applicability	If high concerns regarding applicability for at least one domain, the prediction model evaluation is judged to have **high concerns regarding applicability**.
Unclear concerns regarding applicability	If unclear concerns (but no "high concern") regarding applicability for at least one domain, the prediction model evaluation is judged to have **unclear concerns regarding applicability** overall.

Overall judgement about risk of bias and applicability of the prediction model evaluation		
Overall judgement of risk of bias	**RISK:** *(low/ high/ unclear)*	**H**
Summary of sources of potential bias: Unrepresentative study sample with participants with no post-stroke dependency or depression, yet deemed likely to experience PSCI without clear indication of how this was decided, outcome assessed with use of MoCA for some participants but not for others' and most likely with knowledge of predictor information, issues with analysis from sample size, to selection of predictors, dealing with missing data, to scoring system.		
Overall judgement of applicability	**CONCERN:** *(low/ high/ unclear)*	**L**
Summary of applicability concerns: Due to selection criteria, it is likely that there will be issues with model generalisability, however the study population nonetheless matches the review question.		

Lin et al., 2003

Step 2: Classify the type of prediction model evaluation

Use the following table to classify the evaluation as model development, model validation or model update, or combination. Different signalling questions apply for different types of prediction model evaluation. If the evaluation does not fit one of these classifications then PROBAST should not be used.

Classify the evaluation based on its aim			
Type of prediction study	**PROBAST boxes to complete**	**Tick as appropriate**	**Definition for type of prediction model study**
Development only	Development	✓	Prediction model development without external validation. These studies may include internal validation methods, such as bootstrapping and cross-validation techniques.
Development and validation	Development and validation	X	Prediction model development combined with external validation in other participants in the same article.
Validation only	Validation	X	External validation of existing (previously developed) model in other participants.

This table should be completed once for each publication being assessed and for each relevant outcome in your review.

Publication reference	Lin, L.-H. et al. Prediction of post-stroke dementia. 2003
Models of interest	Formula based on demographic, clinical and performance-based variables
Outcome of interest	Poststroke dementia

Step 3: Assess risk of bias and applicability

PROBAST is structured as four key domains. Each domain is judged for risk of bias (low, high or unclear) and includes signalling questions to help make judgements. Signalling questions are rated as yes (Y), probably yes (PY), probably no (PN), no (N) or no information (NI). All signalling questions are phrased so that "yes" indicates absence of bias. Any signalling question rated as "no" or "probably no" flags the potential for bias; you will need to use your judgement to determine whether the domain should be rated as "high", "low" or "unclear" risk of bias. The guidance document contains further instructions and examples on rating signalling questions and risk of bias for each domain. The first three domains are also rated for concerns regarding applicability (low/ high/ unclear) to your review question defined above.
Complete all domains separately for each evaluation of a distinct model. Shaded boxes indicate where signalling questions do not apply and should not be answered.

DOMAIN 1: Participants			
A. Risk of Bias			
Describe the sources of data and criteria for participant selection: Prospective cohort. "All of the participants were recruited from patients with acute ischemic stroke admitted consecutively to the neurology department from November 1, 1995, to October 31, 1999. Stroke was defined as a rapidly developing clinical manifestation of a focal loss of cerebral function lasting >24 hours. Diagnosis was made by neurologists based on clinical symptoms and confirmed by the findings of neuroimaging studies. Patients with prior cerebrovascular events were also included, whereas those with TIA were excluded. Additional exclusions were those associated with other primary brain lesions (e.g., trauma, tumor, and Parkinson's disease) or severe medical comorbidity (e.g., terminal cancer). To investigate the clinical determinants of new-onset dementia 3 months after stroke onset, we also excluded patients with Alzheimer's disease and possible dementia resulting from other nonvascular etiologies."			
		Dev	Val
1.9 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?		Y	
1.10 Were all inclusions and exclusions of participants appropriate?		Y	
Risk of bias introduced by selection of participants	**RISK:** *(low/ high/ unclear)*	**L**	

Rationale of bias rating:

Exclusion criteria seem appropriate as future cognitive outcomes are less of a priority where patients condition may be terminal and are less likely to be assessed in a real-world healthcare setting, cognitive impairment due to other neurological conditions would confound results, and inclusion of patients with dementia may have led to overoptimistic model estimations.

B. Applicability

Describe included participants, setting and dates:

"Kaohsiung Medical University Hospital is one of the largest medical centers in southern Taiwan, with 23 medical departments and 1,251 beds. It serves the citizens of Kaohsiung City (approximately 1.5 million) as well as people throughout southern Taiwan. Most stroke patients are admitted through the emergency department and then referred to departments such as internal medicine, neurology, or neurosurgery wards for subsequent care.
All of the participants were recruited from patients with acute ischemic stroke admitted consecutively to the neurology department from November 1, 1995, to October 31, 1999."

Concern that the included participants and setting do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	

Rationale of applicability rating:

Participants and setting match the review question and the study sample seems representative of stroke population of interest, although applicability may be affected by cultural, geographical and healthcare settings and recruitment time (started 25 years ago).

DOMAIN 2: Predictors

A. Risk of Bias

List and describe predictors included in the final model, e.g. definition and timing of assessment:

Age of 65 years or older, previous occupation as a labourer, prior stroke, left carotid vascular territory, moderate to severe stroke severity, cognitive impairment, poorer functional status at admission.
"A structured medical history was obtained, and neurologic, functional, cognitive, and neuroimaging examinations were performed within 7 to 10 days of hospitalization. Stroke severity was assessed with the NIH Stroke Scale (NIHSS) administered by the same specially trained nurse. A culturally adapted version of the Mini-Mental State Examination (MMSE) to assess present cognitive status was performed by a specially trained psychologist. Functional status was assessed based on the ability of subjects to perform the motor items of the Functional Independence Measure (FIM) instrument by two senior physical therapists trained in using this instrument. The neurologists and neuroradiologists recorded the MRI findings according to the methods of our previous study for the infarct location (cortical or subcortical area), vascular lesion territory (left or right carotid or vertebrobasilar artery), single or multiple vascular lesion, and stroke mechanism (lacunar, thrombotic, and embolic). Major medical data, prior stroke, previous cardiovascular diseases (e.g., hypertension, diabetes, heart diseases, and hypercholesterolemia), and medical complications were collected from medical records within 1 month after discharge."

		Dev	Val
2.13 Were predictors defined and assessed in a similar way for all participants?		Y	
2.14 Were predictor assessments made without knowledge of outcome data?		Y	
2.15 Are all predictors available at the time the model is intended to be used?		PY	
Risk of bias introduced by predictors or their assessment	**RISK:** *(low/ high/ unclear)*	**L**	
Rationale of bias rating: Data on predictors collected by qualified examiners using recognised and validated tools prior to outcome assessment.			
B. Applicability			
Concern that the definition, assessment or timing of predictors in the model do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating: All aspects related to predictors match the review question, chosen assessment tools are widely accessible and used.			

DOMAIN 3: Outcome
A. Risk of Bias
Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination: "All patients who had survived at 3 months after the stroke were scheduled to undergo follow-up. A version of the neuropsychological test battery adapted for use in the Taiwanese population was administered by psychologists along with the Hamilton Depression Rating Scale (HDRS) as a screen for significant mood disorder and the severity rating of current depression. This battery included the following measures: MMSE orientation items; verbal memory (the Selective Reminding Test); language ability including naming (selected items from the Boston Naming Test), verbal fluency (letter and category fluency subtest of the Boston Diagnostic Aphasia Examination [BDAE]), comprehension (the complex ideation subtest of the BDAE), and repetition (selected items from the repetition subtest of the BDAE); visuospatial ability (items selected from the Rosen Drawing Test and the matching task of the Benton Visual Retention Test); abstractive reasoning (the similarities subtest of the Wechsler Adult Intelligence Scale–Revised and the nonverbal identities and oddities subtest of the Mattis Dementia Rating Scale); and attention function (cancellation tasks using shapes and letters as targets). Senior neurologists collected detailed clinical data, performed the neurobehavioral examinations, and scored the Hachinski Ischemia Scale and the Clinical Dementia Rating Scale (CDR). Furthermore, the NIHSS and the FIM motor scale were reassessed." "Clinical diagnosis of dementia of various types or other mental disorders was made by senior neurologists and psychologists based on the findings of detailed clinical and cognitive evaluations of the patients. Diagnosis of dementia made 3 months after the stroke was based on the Neurologic

Adaptation of the 10th edition of the International Classification of Disease (ICD-10NA) criteria, which require impairment in memory and at least two other cognitive domains and which should be severe enough to affect activities of daily living. VaD was defined as new onset of dementia after stroke using the criteria of the National Institute of Neurologic Disorders and Stroke/Association internationale pour la recherche et l'enseignement en neurosciences. AD was diagnosed according to the criteria of the National Institute of Neurologic and Communicative Disorders and Stroke/Alzheimer's Disease and Related Disorders Association. Other subtypes of dementia were diagnosed based on the CERAD dementia assessment package. All available data were reviewed at a consensus conference to establish a diagnosis. The interrater reliability between two behavioral neurologists involved in the collection and assessment of diagnostic data on neurologic and cognitive function was high."

		Dev	Val
3.25 Was the outcome determined appropriately?		Y	
3.26 Was a pre-specified or standard outcome definition used?		Y	
3.27 Were predictors excluded from the outcome definition?		PN	
3.28 Was the outcome defined and determined in a similar way for all participants?		Y	
3.29 Was the outcome determined without knowledge of predictor information?		PN	
3.30 Was the time interval between predictor assessment and outcome determination appropriate?		Y	
Risk of bias introduced by the outcome or its determination	**RISK:** *(low/ high/ unclear)*	**H**	

Rationale of bias rating:

Outcome seems to have been determined appropriately in terms of tools and criteria used, as well as level of examiner expertise, however a consensus meeting and taking into account "all available data", indicates that predictors were considered when making the diagnosis, which introduces risk of overestimating the strength of association between predictors and outcome.

B. Applicability

At what time point was the outcome determined:

3 months post-stroke

If a composite outcome was used, describe the relative frequency/distribution of each contributing outcome:

N/A

Concern that the outcome, its definition, timing or determination do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	

Rationale of applicability rating:

All aspect of the outcome match the review question.

DOMAIN 4: Analysis
Risk of Bias
Describe numbers of participants, number of candidate predictors, outcome events and events per candidate predictor: 283 participants, 28 candidate predictors, 26 outcome events, EPV = 0.93
Describe how the model was developed (for example in regards to modelling technique (e.g. survival or logistic modelling), predictor selection, and risk group definition): "Patients were divided into two groups according to the diagnosis of dementia. The sociodemographic and clinical data at baseline and 3 months after stroke were compared between these two groups. Education was dichotomized into low (<=6 years) and high (>6 years). Economic status was grouped into low (housewife, unskilled, semiskilled, and skilled trade or craft) and high (clerical/office worker, manager business/government, and professional/technical) social class levels. Previous primary occupation was classified into manual labor and nonmanual labor categories. The severity of stroke was classified as mild (NIHSS 0 to 7) and moderate to severe (NIHSS > 7). Based on our previous study using MMSE to identify subjects with cognitive impairment, different cutoff points for three educational levels were used: <16 for illiterate, <21 for grade school literate, and <24 for junior high school and higher education literate. The functional status by FIM motor score at admission was divided into poor to middle (score 13 to 47) and high (score 48 to 91) score categories. Univariate statistical analysis (chi2 test or Fisher's exact test as appropriate) was carried out to examine the associations between the demographic and clinical factors of stroke patients and the development of PSD. The factors identified as significant by univariate analysis were then applied as prospective predictors to construct a model for predicting the likelihood of PSD. Stepwise logistic regression was employed to construct a predictive model. Colinearity among potential predictors was evaluated using the Spearman rank correlation. Variables with moderate to high intercorrelations (Spearman rank correlation coefficient *r*s >=0.50 or *r*s <=-0.50) were regarded as colinear and consequently not entered together in the same regression analysis. All possible combinations were examined to find the best model. The final model was selected based on goodness of fit."
Describe whether and how the model was validated, either internally (e.g. bootstrapping, cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different setting, different type of participants): Apparent validation only
Describe the performance measures of the model, e.g. (re)calibration, discrimination, (re)classification, net benefit, and whether they were adjusted for optimism: Classification measure only.
Describe any participants who were excluded from the analysis: 69 patients were not interviewed (including 46 who had moved or lost contact or refused follow-up, 10 who had incomplete data collection, 8 who had died prior to the follow-up at 3 months, and 5 who lived in long-term care institutions). Patients who were not examined 3 months after stroke had scores showing significantly more impairment than patients who were examined on the NIHSS and the FIM motor score administered within 7 to 10 days of hospitalization.

Describe missing data on predictors and outcomes as well as methods used for missing data: Participants with missing data were excluded from analysis.			
		Dev	Val
4.37 Were there a reasonable number of participants with the outcome?		N	
4.38 Were continuous and categorical predictors handled appropriately?		PN	
4.39 Were all enrolled participants included in the analysis?		N	
4.40 Were participants with missing data handled appropriately?		N	
4.41 Was selection of predictors based on univariable analysis avoided?		N	
4.42 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?		N	
4.43 Were relevant model performance measures evaluated appropriately?		N	
4.44 Were model overfitting and optimism in model performance accounted for?		N	
4.45 Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?		Y	
Risk of bias introduced by the analysis	**RISK:** *(low/ high/ unclear)*	**H**	
Rationale of bias rating: Too few participants relative to candidate predictors, continuous predictors were dichotomised with atypical choice of cut-off for NIHSS (usually moderate stroke from 5 points onwards), predictors selected based on univariable associations, participants with missing data and no follow-up excluded from the analysis, but shown to have more severe strokes than included patients, no measures of discrimination or calibration, only apparent validation.			

Step 4: Overall assessment

Use the following tables to reach overall judgements about risk of bias and concerns regarding applicability of the prediction model evaluation (development and/or validation) across all assessed domains.
Complete for each evaluation of a distinct model.

Reaching an overall judgement about risk of bias of the prediction model evaluation	
Low risk of bias	If all domains were rated low risk of bias. If a prediction model was developed without any external validation, and it was rated as low risk of bias for all domains, consider downgrading to **high risk of bias**. Such a model can only be considered as low risk of bias, if the development was based on a very large data set and included some form of internal validation.
High risk of bias	If at least one domain is judged to be at **high risk of bias**.

Unclear risk of bias	If an unclear risk of bias was noted in at least one domain and it was low risk for all other domains.

Reaching an overall judgement about applicability of the prediction model evaluation	
Low concerns regarding applicability	If low concerns regarding applicability for all domains, the prediction model evaluation is judged to have **low concerns regarding applicability**.
High concerns regarding applicability	If high concerns regarding applicability for at least one domain, the prediction model evaluation is judged to have **high concerns regarding applicability.**
Unclear concerns regarding applicability	If unclear concerns (but no "high concern") regarding applicability for at least one domain, the prediction model evaluation is judged to have **unclear concerns regarding applicability** overall.

Overall judgement about risk of bias and applicability of the prediction model evaluation		
Overall judgement of risk of bias	**RISK:** *(low/ high/ unclear)*	**H**
Summary of sources of potential bias: Diagnosis of outcome most likely involving knowledge of predictors, multiple issues with analysis from sample size, through predictor categorisation and selection, dealing with missing data, to performance estimation and lack of appropriate model validation.		
Overall judgement of applicability	**CONCERN:** *(low/ high/ unclear)*	**L**
Summary of applicability concerns: Participants, predictors and outcome all match review question and population, setting and information of interest.		

Munsch et al., 2016

Step 2: Classify the type of prediction model evaluation

Use the following table to classify the evaluation as model development, model validation or model update, or combination. Different signalling questions apply for different types of prediction model evaluation. If the evaluation does not fit one of these classifications then PROBAST should not be used.

Classify the evaluation based on its aim			
Type of prediction study	**PROBAST boxes to complete**	**Tick as appropriate**	**Definition for type of prediction model study**
Development only	Development	✓	Prediction model development without external validation. These studies may include internal validation methods, such as bootstrapping and cross-validation techniques.
Development and validation	Development and validation	X	Prediction model development combined with external validation in other participants in the same article.
Validation only	Validation	X	External validation of existing (previously developed) model in other participants.

This table should be completed once for each publication being assessed and for each relevant outcome in your review.

Publication reference	Munsch, F. Stroke location is an independent predictor of cognitive outcome. 2016
Models of interest	Formula based on demographic, clinical and imaging variables
Outcome of interest	Good post-stroke cognitive outcome

Step 3: Assess risk of bias and applicability

PROBAST is structured as four key domains. Each domain is judged for risk of bias (low, high or unclear) and includes signalling questions to help make judgements. Signalling questions are rated as yes (Y), probably yes (PY), probably no (PN), no (N) or no information (NI). All signalling questions are phrased so that "yes" indicates absence of bias. Any signalling question rated as "no" or "probably no" flags the potential for bias; you will need to use your judgement to determine whether the domain should be rated as "high", "low" or "unclear" risk of bias. The guidance document contains further instructions and examples on rating signalling questions and risk of bias for each domain. The first three domains are also rated for concerns regarding applicability (low/ high/ unclear) to your review question defined above.

Complete all domains separately for each evaluation of a distinct model. Shaded boxes indicate where signalling questions do not apply and should not be answered.

DOMAIN 1: Participants			
A. Risk of Bias			
Describe the sources of data and criteria for participant selection: Prospective cohort; "Primary inclusion criteria were men and women, older than 18 years old, with a clinical diagnosis of minor-to-severe supratentorial cerebral infarct (NIHSS between 1 and 25) between 24 and 72 hours after the onset. Exclusion criteria were history of symptomatic cerebral infarct with functional deficit (prestroke modified Rankin Scale [mRS] score ≥1), infratentorial stroke, history of severe cognitive impairment (dementia), or psychiatric troubles matching to axis 1 of the Diagnostic and Statistical Manual of Mental Disorders, Fourth edition (DSM-4) criteria except for major depression, coma, pregnant or breast-feeding women, and contraindications to MRI."			
		Dev	Val
1.11 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?		Y	
1.12 Were all inclusions and exclusions of participants appropriate?		Y	
Risk of bias introduced by selection of participants	**RISK:** *(low/ high/ unclear)*	**L**	
Rationale of bias rating: inclusion of patients with dementia may have led to overoptimistic model estimations, as the study also aimed to predict functional outcome, inclusion of participants with previous functional disability due to stroke may have caused a similar issue, other exclusion criteria based on practical considerations, unclear why stroke severity was limited to a score of 25 – this may be just the max in the sample, however, this is unlikely to introduce high risk of bias, as very severe stroke are relatively uncommon now and in such cases prediction of longer-term cognitive outcome is not a high priority.			
B. Applicability			
Describe included participants, setting and dates: Most likely conducted in France, 428 consecutive patients presenting a suspected supratentorial ischemic stroke from June 2012 to February 2015.			
Concern that the included participants and setting do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating: Information regarding study setting is limited, however review question is very broad so it is unlikely that it would not match.			

DOMAIN 2: Predictors

A. Risk of Bias

List and describe predictors included in the final model, e.g. definition and timing of assessment:

NIHSS (continuous variable), Age (continuous variable), Lesion volume (continuous in cm3), Log of stroke location (continuous in voxels).
At baseline, the NIHSS was recorded between 24 and 72 hours after stroke onset, at the time of the MRI scan.
"We used the VLSM method implemented in the nonparametric mapping toolbox included in the MRIcron software package (MRIcron, Verion 6.6.2013). This method establishes a relationship between the presence or lack of a lesion and a behavioral score on a voxel-by-voxel basis. For each voxel, a Brunner–Munzel rank order test was performed to determine whether the behavioral score is significantly different between the lesioned and nonlesioned group. We built maps of functional and cognitive eloquent regions using, respectively, mRS and MoCA measured at 3 months as behavioral scores. A subanalysis was conducted on a short MoCA (sMOCA) in which the items naming and language have been removed. The resulting Z score maps were controlled for multiple comparisons using the false discovery rate correction to ensure a false-positive rate of $P<0.05$. The eloquent regions were identified using the Automated Anatomic Labeling, Brodmann, and JHU-WhiteMatterlabels-1mm atlases available in the MRIcron software package. The objective was to use the VLSM maps, which showed the eloquent areas in terms of mRS and MoCA scores, to predict, respectively, the functional and cognitive outcomes at 3 months for a new stroke patient. For that purpose, we overlapped the patient's lesion binary mask on each VLSM map. Then, we extracted all significant Z scores (corresponding to eloquent voxels that survived a 5% false discovery rate cutoff threshold) contained in the lesion, using a home-made program developed in Matlab (Mathworks Natick, Massachusetts). Finally, using the R software package (Version 3.0.1), we calculated the number of eloquent voxels. This quantitative variable contains the information of location and will be referred to as stroke location in the following sections."

		Dev	Val
2.16 Were predictors defined and assessed in a similar way for all participants?		Y	
2.17 Were predictor assessments made without knowledge of outcome data?		N	
2.18 Are all predictors available at the time the model is intended to be used?		N	
Risk of bias introduced by predictors or their assessment	**RISK:** *(low/ high/ unclear)*	**H**	

Rationale of bias rating:

Voxels were identified based on outcome data from the same participants who were involved in developing the prognostic model and on whom the model's performance was assessed – high risk of overestimating the strength of association between predictors and outcome.

B. Applicability

Concern that the definition, assessment or timing of predictors in the model do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**H**	

Rationale of applicability rating:

To estimate stroke location using this method required use of outcome data, the review question specified that information on predictors should be collected prior to the occurrence of the outcome; it is unclear how an independent researcher/clinician could apply the method of voxel estimation themselves – the method does not appear easily accessible in the setting it would potentially be applied in.

DOMAIN 3: Outcome

A. Risk of Bias

Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination:

At 3-month follow-up, all patients underwent a standardized battery of clinical testing including, among others, … the Montreal Cognitive Assessment (MoCA) to assess cognitive deficits.
A favorable cognitive outcome was defined as a MoCA score >25.

		Dev	Val
3.31 Was the outcome determined appropriately?		Y	
3.32 Was a pre-specified or standard outcome definition used?		Y	
3.33 Were predictors excluded from the outcome definition?		Y	
3.34 Was the outcome defined and determined in a similar way for all participants?		PY	
3.35 Was the outcome determined without knowledge of predictor information?		NI	
3.36 Was the time interval between predictor assessment and outcome determination appropriate?		Y	
Risk of bias introduced by the outcome or its determination	**RISK:** *(low/ high/ unclear)*	**L**	

Rationale of bias rating:

It is unclear whether examiners were blinded to predictor information and test scores may be to a certain extent influenced by examiner behaviour, yet given the binary nature of the outcome, this would likely only affect participants with actual scores near the cut off. It is also unclear who carried out the assessment, however administering the MoCA does not require extensive and/or specialist training.

B. Applicability

At what time point was the outcome determined:

3 months post-stroke

If a composite outcome was used, describe the relative frequency/distribution of each contributing outcome:

N/A

Concern that the outcome, its definition, timing or determination do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	

Rationale of applicability rating:

All aspects of the outcome match the review question.

DOMAIN 4: Analysis			
Risk of Bias			
Describe numbers of participants, number of candidate predictors, outcome events and events per candidate predictor: 198 participants, 4 candidate predictors, 77 outcome events, EPV = 19.3			
Describe how the model was developed (for example in regards to modelling technique (e.g. survival or logistic modelling), predictor selection, and risk group definition): Logistic regression, predictors specified before analysis with 2 candidate models later compared against one another in terms of discrimination, all predictors were left continuous			
Describe whether and how the model was validated, either internally (e.g. bootstrapping, cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different setting, different type of participants): Internal model validation using 10 fold-cross validation and sample splitting based on recruitment period.			
Describe the performance measures of the model, e.g. (re)calibration, discrimination, (re)classification, net benefit, and whether they were adjusted for optimism: Discrimination			
Describe any participants who were excluded from the analysis: 28 excluded due to missing MRI sequences, 23 due to images with insufficient quality, 26 lost to follow-up, 13 died, 10 with missing outcome data			
Describe missing data on predictors and outcomes as well as methods used for missing data: Participants with missing data were excluded from the analysis.			
		Dev	Val
4.46 Were there a reasonable number of participants with the outcome?		PY	
4.47 Were continuous and categorical predictors handled appropriately?		Y	
4.48 Were all enrolled participants included in the analysis?		N	
4.49 Were participants with missing data handled appropriately?		N	
4.50 Was selection of predictors based on univariable analysis avoided?		Y	
4.51 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?		N	
4.52 Were relevant model performance measures evaluated appropriately?		N	
4.53 Were model overfitting and optimism in model performance accounted for?		Y	
4.54 Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?		Y	
Risk of bias introduced by the analysis	**RISK:** *(low/ high/ unclear)*	**H**	

Rationale of bias rating: Data of less than half the enrolled participants (428) were use for model development, among other reasons, with some participants assigned to a validation set, and participants with missing data or lost to follow-up being excluded from the analysis; also no assessment of calibration.

Step 4: Overall assessment

Use the following tables to reach overall judgements about risk of bias and concerns regarding applicability of the prediction model evaluation (development and/or validation) across all assessed domains.
Complete for each evaluation of a distinct model.

Reaching an overall judgement about risk of bias of the prediction model evaluation	
Low risk of bias	If all domains were rated low risk of bias. If a prediction model was developed without any external validation, and it was rated as low risk of bias for all domains, consider downgrading to **high risk of bias**. Such a model can only be considered as low risk of bias, if the development was based on a very large data set and included some form of internal validation.
High risk of bias	If at least one domain is judged to be at **high risk of bias**.
Unclear risk of bias	If an unclear risk of bias was noted in at least one domain and it was low risk for all other domains.

Reaching an overall judgement about applicability of the prediction model evaluation	
Low concerns regarding applicability	If low concerns regarding applicability for all domains, the prediction model evaluation is judged to have **low concerns regarding applicability**.
High concerns regarding applicability	If high concerns regarding applicability for at least one domain, the prediction model evaluation is judged to have **high concerns regarding applicability**.
Unclear concerns regarding applicability	If unclear concerns (but no "high concern") regarding applicability for at least one domain, the prediction model evaluation is judged to have **unclear concerns regarding applicability** overall.

Overall judgement about risk of bias and applicability of the prediction model evaluation		
Overall judgement of risk of bias	**RISK:** *(low/ high/ unclear)*	**H**
Summary of sources of potential bias: Issues relating to one of the main predictors of interest being determined based on outcome data and exclusion of participants from the analysis due to missing data and loss to follow-up, model calibration was not assessed.		
Overall judgement of applicability	**CONCERN:** *(low/ high/ unclear)*	**H**

Summary of applicability concerns:
Review question specified that information on predictors ought to be known prior to outcome occurrence, also the method for estimating stroke location would not be accessible in a clinical setting and would likely even be challenging to apply in a research setting.

Salihovic et al., 2018

Step 2: Classify the type of prediction model evaluation

Use the following table to classify the evaluation as model development, model validation or model update, or combination. Different signalling questions apply for different types of prediction model evaluation. If the evaluation does not fit one of these classifications then PROBAST should not be used.

Classify the evaluation based on its aim			
Type of prediction study	**PROBAST boxes to complete**	**Tick as appropriate**	**Definition for type of prediction model study**
Development only	Development	✓	Prediction model development without external validation. These studies may include internal validation methods, such as bootstrapping and cross-validation techniques.
Development and validation	Development and validation	X	Prediction model development combined with external validation in other participants in the same article.
Validation only	Validation	X	External validation of existing (previously developed) model in other participants.

This table should be completed once for each publication being assessed and for each relevant outcome in your review.

Publication reference	Salihovic, D. Cognitive syndromes after the first stroke. 2018
Models of interest	Decision tree based on cognitive performance
Outcome of interest	Dementia/ vascular cognitive syndromes

Step 3: Assess risk of bias and applicability

PROBAST is structured as four key domains. Each domain is judged for risk of bias (low, high or unclear) and includes signalling questions to help make judgements. Signalling questions are rated as yes (Y), probably yes (PY), probably no (PN), no (N) or no information (NI). All signalling questions are

phrased so that “yes” indicates absence of bias. Any signalling question rated as “no” or “probably no” flags the potential for bias; you will need to use your judgement to determine whether the domain should be rated as “high”, “low” or “unclear” risk of bias. The guidance document contains further instructions and examples on rating signalling questions and risk of bias for each domain. The first three domains are also rated for concerns regarding applicability (low/ high/ unclear) to your review question defined above.
Complete all domains separately for each evaluation of a distinct model. Shaded boxes indicate where signalling questions do not apply and should not be answered.

DOMAIN 1: Participants			
A. Risk of Bias			
Describe the sources of data and criteria for participant selection: “The study was prospective and included 275 patients with acute stroke (ischemic and hemorrhagic), both sexes and all age groups, who were hospitalized at the Department of Neurology, University Clinical Center Tuzla (from September 1, 2011, to August 31, 2012.). Excluding criteria were subarachnoid hemorrhage, recurrent stroke, or mortality in the first 3 months after the patient was included in the study, existence of cognitive impairment before the beginning of the study (based on medical records). Also, study did not include patients whose bad general somatic state enabled quality testing, as well as patients with aphasia who could not do all tests.”			
		Dev	Val
1.13 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?		Y	
1.14 Were all inclusions and exclusions of participants appropriate?		Y	
Risk of bias introduced by selection of participants	**RISK:** *(low/ high/ unclear)*	**L**	
Rationale of bias rating: Inclusion of patients with cognitive impairment may lead to overoptimistic model estimations; recurrent stroke is likely to affect outcome, yet difficult to account for if using baseline predictors; other exclusion criteria based on practical considerations; main concern relates to excluding patients who died within the first 3 months, yet this will be addressed in the analysis domain			
B. Applicability			
Describe included participants, setting and dates: Patients with acute stroke, hospitalized at the Department of Neurology, University Clinical Center Tuzla (September 1, 2011 to August 31, 2012)			
Concern that the included participants and setting do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating: Participants and setting match review question			

DOMAIN 2: Predictors			
A. Risk of Bias			
List and describe predictors included in the final model, e.g. definition and timing of assessment: "Patients were tested in three occasions; first testing was after 3 months, second testing after 6 months, and third after 12 months of stroke. For the evaluation of cognitive functions, the following measure scales were used: Mini-Mental Status Examination MMSE, Montreal Cognitive Assessment, Wechsler's Intelligence Scale WB-II, Wechsler's scale of memory, Rey-Osterrieth complex figure test—RCFT"			
		Dev	Val
2.19 Were predictors defined and assessed in a similar way for all participants?		PY	
2.20 Were predictor assessments made without knowledge of outcome data?		PY	
2.21 Are all predictors available at the time the model is intended to be used?		Y	
Risk of bias introduced by predictors or their assessment	**RISK:** *(low/ high/ unclear)*	**L**	
Rationale of bias rating: Unclear description, however seems that standard tests were used for assessment of cognitive performance (predictors), and with a prospective study design it can be assumed that predictors were estimated prior to the outcome (and therefore without knowledge of if).			
B. Applicability			
Concern that the definition, assessment or timing of predictors in the model do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating: Unclear description, yet review question is very broad and there is no indication that chosen predictors do not match it			

DOMAIN 3: Outcome		
A. Risk of Bias		
Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination: Patients were divided into following groups: dementia of strategic infarct (DSI), cortical dementia (CD), sub cortical dementia (SCD), hemorrhagic dementia (HD), and patients without dementia. The VCI diagnosis was based on diagnostic criteria (DSM-4 and ICD-10), clinical exams, and neuropsychological testing.		
	Dev	Val
3.37 Was the outcome determined appropriately?	Y	
3.38 Was a pre-specified or standard outcome definition used?	NI	
3.39 Were predictors excluded from the outcome definition?	PN	
3.40 Was the outcome defined and determined in a similar way for all participants?	PY	

3.41 Was the outcome determined without knowledge of predictor information?		PN	
3.42 Was the time interval between predictor assessment and outcome determination appropriate?		NI	
Risk of bias introduced by the outcome or its determination	**RISK:** *(low/ high/ unclear)*	**H**	
Rationale of bias rating: Given criteria for VCI, it seems that performance on cognitive tests would have been taken into account to make a diagnosis of dementia.			
B. Applicability			
At what time point was the outcome determined: 12 months post-stroke *If a composite outcome was used, describe the relative frequency/distribution of each contributing outcome:* N/A			
Concern that the outcome, its definition, timing or determination do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating: Outcome definition, timing and assessment match review question.			

DOMAIN 4: Analysis
Risk of Bias
Describe numbers of participants, number of candidate predictors, outcome events and events per candidate predictor: 275 participants, unclear number of candidate predictors (min. 3), 85 participants in smaller outcome group, EPV unknown.
Describe how the model was developed (for example in regards to modelling technique (e.g. survival or logistic modelling), predictor selection, and risk group definition): A CHAID algorithm analysis was used to develop the decision tree models. CHAID decision trees are nonparametric procedures that make no assumptions of the underlying data. This algorithm determines how continuous and/or categorical independent variables best combine to predict a binary outcome based on "if-then" logic by portioning each independent variable into mutually exclusive subsets based on homogeneity of the data.
Describe whether and how the model was validated, either internally (e.g. bootstrapping, cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different setting, different type of participants): No report on validation.
Describe the performance measures of the model, e.g. (re)calibration, discrimination, (re)classification, net benefit, and whether they were adjusted for optimism: No performance measures reported

Describe any participants who were excluded from the analysis: Patients were excluded if they died during the first 3 months after stroke.			
Describe missing data on predictors and outcomes as well as methods used for missing data: No mention of missing data.			
		Dev	Val
4.55 Were there a reasonable number of participants with the outcome?		PN	
4.56 Were continuous and categorical predictors handled appropriately?		NI	
4.57 Were all enrolled participants included in the analysis?		N	
4.58 Were participants with missing data handled appropriately?		N	
4.59 Was selection of predictors based on univariable analysis avoided?		PY	
4.60 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?		N	
4.61 Were relevant model performance measures evaluated appropriately?		N	
4.62 Were model overfitting and optimism in model performance accounted for?		N	
4.63 Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?		N/A	
Risk of bias introduced by the analysis	**RISK:** *(low/ high/ unclear)*	**H**	
Rationale of bias rating: CHAID analysis requires high sample size, because of the data splitting it entails; participants who died within 3 months of index stroke were excluded; no mention of how missing data were handled; CHAID is considered an exploratory method; no mention of any form of validation.			

Step 4: Overall assessment

Use the following tables to reach overall judgements about risk of bias and concerns regarding applicability of the prediction model evaluation (development and/or validation) across all assessed domains.
Complete for each evaluation of a distinct model.

Reaching an overall judgement about risk of bias of the prediction model evaluation	
Low risk of bias	If all domains were rated low risk of bias. If a prediction model was developed without any external validation, and it was rated as low risk of bias for all domains, consider downgrading to **high risk of bias**. Such a model can only be considered as low risk of bias, if the development was based on a very large data set and included some form of internal validation.
High risk of bias	If at least one domain is judged to be at **high risk of bias**.

Unclear risk of bias	If an unclear risk of bias was noted in at least one domain and it was low risk for all other domains.

Reaching an overall judgement about applicability of the prediction model evaluation	
Low concerns regarding applicability	If low concerns regarding applicability for all domains, the prediction model evaluation is judged to have **low concerns regarding applicability**.
High concerns regarding applicability	If high concerns regarding applicability for at least one domain, the prediction model evaluation is judged to have **high concerns regarding applicability**.
Unclear concerns regarding applicability	If unclear concerns (but no "high concern") regarding applicability for at least one domain, the prediction model evaluation is judged to have **unclear concerns regarding applicability** overall.

Overall judgement about risk of bias and applicability of the prediction model evaluation		
Overall judgement of risk of bias	**RISK:** *(low/ high/ unclear)*	**H**
Summary of sources of potential bias: Main issues relate to outcome likely being based on predictor information and analysis limitations- small sample size, excluding patients who died, no form performance estimation/validation.		
Overall judgement of applicability	**CONCERN:** *(low/ high/ unclear)*	**L**
Summary of applicability concerns: All assessed aspects match review question.		

Kostalova et al., 2012

Step 2: Classify the type of prediction model evaluation

Use the following table to classify the evaluation as model development, model validation or model update, or combination. Different signalling questions apply for different types of prediction model evaluation. If the evaluation does not fit one of these classifications then PROBAST should not be used.

Classify the evaluation based on its aim			
Type of prediction study	**PROBAST boxes to complete**	**Tick as appropriate**	**Definition for type of prediction model study**
Development only	Development	✓	Prediction model development without external validation. These studies may include internal validation methods, such as bootstrapping and cross-validation techniques.

Development and validation	Development and validation	X	Prediction model development combined with external validation in other participants in the same article.
Validation only	Validation	X	External validation of existing (previously developed) model in other participants.

This table should be completed once for each publication being assessed and for each relevant outcome in your review.

Publication reference	Kostalova, M. et al. Towards a predictive model for post-stroke delirium. 2012
Models of interest	Preliminary model based on demographic, clinical and laboratory variables
Outcome of interest	Post-stroke delirium

Step 3: Assess risk of bias and applicability

PROBAST is structured as four key domains. Each domain is judged for risk of bias (low, high or unclear) and includes signalling questions to help make judgements. Signalling questions are rated as yes (Y), probably yes (PY), probably no (PN), no (N) or no information (NI). All signalling questions are phrased so that "yes" indicates absence of bias. Any signalling question rated as "no" or "probably no" flags the potential for bias; you will need to use your judgement to determine whether the domain should be rated as "high", "low" or "unclear" risk of bias. The guidance document contains further instructions and examples on rating signalling questions and risk of bias for each domain. The first three domains are also rated for concerns regarding applicability (low/ high/ unclear) to your review question defined above.
Complete all domains separately for each evaluation of a distinct model. Shaded boxes indicate where signalling questions do not apply and should not be answered.

DOMAIN 1: Participants
A. Risk of Bias
Describe the sources of data and criteria for participant selection: A prospective observational study in a cohort of consecutive patients with ischemic stroke or intracerebral haematoma admitted within 24 hours of stroke onset. The inclusion criteria were: admission diagnosis of cerebral infarction or intracerebral haemorrhage; that a delirium assessment could be carried out within 24 hours of stroke onset; and approval of the patient or his/her relatives. A priori exclusion criteria were: duration of stroke symptoms and signs < 24 hours, history of severe head trauma or neurosurgery at any time before stroke; subarachnoid haemorrhage, venous infarction or brain tumour; history of psychosis; patients who did not speak Czech; and patients who were comatose or stuporous on admission and did not improve during the first week post-stroke, with a Richmond agitation and sedation scale (RASS) score <=-4.

		Dev	Val
1.15 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?		Y	
1.16 Were all inclusions and exclusions of participants appropriate?		Y	
Risk of bias introduced by selection of participants	**RISK:** *(low/ high/ unclear)*	**L**	
Rationale of bias rating: Selection criteria appropriate as overall are aimed to exclude participants who are not within the focus of the study (not stroke patients), for whom an assessment would not be appropriate due to severity of condition and who may have confounding conditions, leading to spurious results.			
B. Applicability			
Describe included participants, setting and dates: The study population was recruited over a 15-month period (between January 2009 and March 2010) from all stroke patients consecutively admitted to a 6-bed stroke unit, part of the Department of Neurology of the University Hospital.			
Concern that the included participants and setting do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating: Participants and setting match the review question focusing on people diagnosed with any type of stroke.			

DOMAIN 2: Predictors			
A. Risk of Bias			
List and describe predictors included in the final model, e.g. definition and timing of assessment: Model 1: age, GGT > 1.02 ukat, bilirubin > 20, ICH, lesion volume > 40 ccm Model 2: age, ICH, lesion volume > 40 ccm, SOFA-Max > 2 Risk factors evaluated prospectively; laboratory markers were estimated multiple times within the 7 day follow-up period, using the most abnormal value for inclusion in the model. The volume of the infarction was calculated from delayed MRI FLAIR 3-D scans performed 4–6 weeks after onset of stroke using the volumetric semi-automatic segmentation method.			
		Dev	Val
2.22 Were predictors defined and assessed in a similar way for all participants?		Y	
2.23 Were predictor assessments made without knowledge of outcome data?		NI	
2.24 Are all predictors available at the time the model is intended to be used?		N	
Risk of bias introduced by predictors or their assessment	**RISK:** *(low/ high/ unclear)*	**L**	

Rationale of bias rating:

Outcome and many predictors were assessed within the same follow-up period, it is therefore possible that a predictor value was recorded after occurrence of delirium, also it seems likely that the clinicians/researchers involved in the data collection were aware of both patient predictors and outcome, however selected predictors seem objective, with lesion volume assessed using a validated, semi-automated method.

B. Applicability

Concern that the definition, assessment or timing of predictors in the model do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**H**	

Rationale of applicability rating:

The timing of predictors does not match the review question as we specified that predictors need to be recorded/known prior to the occurrence of the outcome.

DOMAIN 3: Outcome

A. Risk of Bias

Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination:

"Evaluation of delirium was performed daily by a panel of specialist, delirium experts working in the University hospital: two neurologists with over 15 years of experience and two neuropsychologists with over 5 years of experience with stroke and intensive care patients, a psychiatrist with over 10 years of experience in psychiatric intensive care and a speech therapist with over 10 years of experience with stroke patients . They applied the criteria for delirium presented in DSM-4 and based their judgements on interviews with family members, information obtained from junior and nursing staff and chart reviews. The core clinical criteria for diagnosis are a disturbance of consciousness with reduced ability to focus, sustain or shift attention; other cognitive or perceptual disturbances; acuity of onset (hours to days) and fluctuation over the course of the day; and presence of an underlying cause, such as a general medical condition, medication, a combination of aetiologies or indeterminate aetiology. With respect to high prevalence of pre-stroke dementia or possibility of disturbances of memory, perception or attention (such as aphasia or neglect) caused by lesions of certain brain regions, the mental status of patients on admission was considered as a new baseline and acute onset and fluctuating course was assessed as positive in the event of new change from this baseline mental status or fluctuation in it. Each evaluation was performed by at least one neurologist and one neuropsychologist; the speech therapist who had performed the introductory logopedic exam was then brought in, together with a psychiatrist if necessary. In parallel with clinical assessment based on DSM-4 criteria, blindly and independently CAM-ICU scoring was used and high sensitivity, specificity and accuracy of this diagnostic instrument was found. The delirium experts standardized their approach to DSM-4 criteria and CAM-ICU scoring over a 3-month training phase via roundtable discussions with attending neurologists and intensive care specialists regarding their approach to standardizing their ICU delirium assessment. The expert evaluation, however, was used as a gold standard both in this methodological study and in current evaluation of risk factors. The first evaluation was made within 24 hours of admission (day 1)

and then daily on 7 consecutive days upon which the patient was accessible to testing irrespective of the result of the first evaluation. Follow-up was stopped in patients who became inaccessible to testing because of worsening of consciousness or death."			
		Dev	Val
3.43 Was the outcome determined appropriately?		Y	
3.44 Was a pre-specified or standard outcome definition used?		Y	
3.45 Were predictors excluded from the outcome definition?		N	
3.46 Was the outcome defined and determined in a similar way for all participants?		Y	
3.47 Was the outcome determined without knowledge of predictor information?		N	
3.48 Was the time interval between predictor assessment and outcome determination appropriate?		N	
Risk of bias introduced by the outcome or its determination	**RISK:** *(low/ high/ unclear)*	**H**	
Rationale of bias rating:			
The method of outcome assessment seems to reach a gold standard, given the applied criteria, undergone training and experts involved, however it appears that predictors were taken into account in this process and therefore the association between predictors and outcome is likely to be overestimated.			
B. Applicability			
At what time point was the outcome determined:			
Within a 7-day period, which was the same as the predictors			
If a composite outcome was used, describe the relative frequency/distribution of each contributing outcome:			
N/A			
Concern that the outcome, its definition, timing or determination do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating:			
The issue with timing relates to the assessment of predictors overlapping with the assessment of the outcome, however in itself the timing of outcome assessment is appropriate given that delirium can develop and recede shortly after admission; all other aspects of the outcome match the review question.			

DOMAIN 4: Analysis
Risk of Bias
Describe numbers of participants, number of candidate predictors, outcome events and events per candidate predictor:
100 participants, 37 candidate predictors, 43 outcome events, EPV = 1.2
Describe how the model was developed (for example in regards to modelling technique (e.g. survival or logistic modelling), predictor selection, and risk group definition):
Where employed, optimal cut-off values of continuous variables were optimized on the basis of ROC analysis. "For multivariate analysis the

backward stepwise logistic regression model was used, entering all the variables with p < 0.10. on univariate analysis. The final set of potential predictive factors and interaction terms (coded as binary variables) was subjected to a stepwise selection algorithm in multivariate logistic regression (driven by maximum likelihood ratio test). Only predictors that were significant and mutually uncorrelated were kept in the final model."

Describe whether and how the model was validated, either internally (e.g. bootstrapping, cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different setting, different type of participants):

"A simple 2-fold cross-validation (randomly assigning database records to two data sets of equal size and using them changeably as training and testing set) was applied."

Describe the performance measures of the model, e.g. (re)calibration, discrimination, (re)classification, net benefit, and whether they were adjusted for optimism:

"Proportion of correctly recognized cases was evaluated to assess the accuracy of the proposed model equations."

Describe any participants who were excluded from the analysis:

19 incomplete protocols - patients who died or became comatose or stuporous during the follow-up.

Describe missing data on predictors and outcomes as well as methods used for missing data:

No information on missing data.

		Dev	Val
4.64 Were there a reasonable number of participants with the outcome?		N	
4.65 Were continuous and categorical predictors handled appropriately?		N	
4.66 Were all enrolled participants included in the analysis?		N	
4.67 Were participants with missing data handled appropriately?		NI	
4.68 Was selection of predictors based on univariable analysis avoided?		N	
4.69 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?		N	
4.70 Were relevant model performance measures evaluated appropriately?		N	
4.71 Were model overfitting and optimism in model performance accounted for?		Y	
4.72 Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?		Y	
Risk of bias introduced by the analysis	**RISK:** *(low/ high/ unclear)*	**H**	

Rationale of bias rating:

Too few participants, which authors indicated by calling this a "pilot attempt", use of data-driven predictor cut-offs, predictor selection based on univariable associations, exclusion of participants lost to follow-up, no estimates of model discrimination or calibration.

Step 4: Overall assessment

Use the following tables to reach overall judgements about risk of bias and concerns regarding applicability of the prediction model evaluation (development and/or validation) across all assessed domains.
Complete for each evaluation of a distinct model.

Reaching an overall judgement about risk of bias of the prediction model evaluation	
Low risk of bias	If all domains were rated low risk of bias. If a prediction model was developed without any external validation, and it was rated as low risk of bias for all domains, consider downgrading to **high risk of bias**. Such a model can only be considered as low risk of bias, if the development was based on a very large data set and included some form of internal validation.
High risk of bias	If at least one domain is judged to be at **high risk of bias**.
Unclear risk of bias	If an unclear risk of bias was noted in at least one domain and it was low risk for all other domains.

Reaching an overall judgement about applicability of the prediction model evaluation	
Low concerns regarding applicability	If low concerns regarding applicability for all domains, the prediction model evaluation is judged to have **low concerns regarding applicability**.
High concerns regarding applicability	If high concerns regarding applicability for at least one domain, the prediction model evaluation is judged to have **high concerns regarding applicability**.
Unclear concerns regarding applicability	If unclear concerns (but no "high concern") regarding applicability for at least one domain, the prediction model evaluation is judged to have **unclear concerns regarding applicability** overall.

Overall judgement about risk of bias and applicability of the prediction model evaluation		
Overall judgement of risk of bias	**RISK:** *(low/ high/ unclear)*	**H**
Summary of sources of potential bias: Main sources relate to inclusion of predictor information in outcome assessment and multiple issues with analysis procedures and estimation of model performance.		
Overall judgement of applicability	**CONCERN:** *(low/ high/ unclear)*	**H**
Summary of applicability concerns: Issues due to assessment of predictor overlapping with outcome assessment, which does not match the review question.		

Kotfis et al., 2019; DELIAS

Step 2: Classify the type of prediction model evaluation

Use the following table to classify the evaluation as model development, model validation or model update, or combination. Different signalling questions apply for different types of prediction model evaluation. If the evaluation does not fit one of these classifications then PROBAST should not be used.

Classify the evaluation based on its aim			
Type of prediction study	**PROBAST boxes to complete**	**Tick as appropriate**	**Definition for type of prediction model study**
Development only	Development	✓	Prediction model development without external validation. These studies may include internal validation methods, such as bootstrapping and cross-validation techniques.
Development and validation	Development and validation	X	Prediction model development combined with external validation in other participants in the same article.
Validation only	Validation	X	External validation of existing (previously developed) model in other participants.

This table should be completed once for each publication being assessed and for each relevant outcome in your review.

Publication reference	Kotfis, K. et al. Could neutrophil-to-lymphocyte ratio (NLR) serve as a potential marker for delirium prediction in patients with acute ischemic stroke? a prospective observational study. 2019
Models of interest	DELIAS score based on clinical and laboratory variables
Outcome of interest	Post-stroke delirium

Step 3: Assess risk of bias and applicability

PROBAST is structured as four key domains. Each domain is judged for risk of bias (low, high or unclear) and includes signalling questions to help make judgements. Signalling questions are rated as yes (Y), probably yes (PY), probably no (PN), no (N) or no information (NI). All signalling questions are phrased so that “yes” indicates absence of bias. Any signalling question rated as “no” or “probably no” flags the potential for bias; you will need to use your judgement to determine whether the domain should be rated as “high”, “low” or “unclear” risk of bias. The guidance document contains further instructions and examples on rating signalling questions and risk of bias for each domain.

The first three domains are also rated for concerns regarding applicability (low/ high/ unclear) to your review question defined above.
Complete all domains separately for each evaluation of a distinct model.
Shaded boxes indicate where signalling questions do not apply and should not be answered.

DOMAIN 1: Participants			
A. Risk of Bias			
Describe the sources of data and criteria for participant selection: Prospective observational study; Consecutive adult patients (age > 18 years) with acute ischemic stroke; "We excluded patients with hematology disorders (5 patients), incomplete laboratory testing (6 patients), or no data regarding follow-up."			
		Dev	Val
1.17 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?		Y	
1.18 Were all inclusions and exclusions of participants appropriate?		Y	
Risk of bias introduced by selection of participants	**RISK:** *(low/ high/ unclear)*	**L**	
Rationale of bias rating: Excluding patients with incomplete data is a potential source of bias, however this will be addressed separately under the analysis domain, other criteria seem appropriate			
B. Applicability			
Describe included participants, setting and dates: Patients admitted to the neurology department of a busy district general hospital in Poland between 30 June 2015 and 31 March 2018. Patients with AIS were admitted to the neurology department within 48 hours of symptom development.			
Concern that the included participants and setting do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating: Both participants and setting match review question.			

DOMAIN 2: Predictors			
A. Risk of Bias			
List and describe predictors included in the final model, e.g. definition and timing of assessment: Hemianopia, aphasia, age, NIHSS on admission, NLR, Leukocytes, CRP The data was collected prospectively by a dedicated member of staff. Blood analysis carried out at admission.			
		Dev	Val
2.25 Were predictors defined and assessed in a similar way for all participants?		NI	

2.26 Were predictor assessments made without knowledge of outcome data?		NI	
2.27 Are all predictors available at the time the model is intended to be used?		NI	
Risk of bias introduced by predictors or their assessment	**RISK:** *(low/ high/ unclear)*	**UN**	
Rationale of bias rating:			
As the primary outcome was assessed within 24 hours, it seems possible that information on some predictors could have been collected after the outcome assessment, which could e.g. affect judgement regarding aphasia, hemianopia and/or NIHSS. Also the method of assessing aphasia and hemianopia is unclear.			
B. Applicability			
Concern that the definition, assessment or timing of predictors in the model do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**UN**	
Rationale of applicability rating:			
Applicability would depend on whether the predictors were all estimated prior to the outcome.			

DOMAIN 3: Outcome			
A. Risk of Bias			
Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination:			
We used the Polish version of the CAM-ICU assessment tool to screen all patients for delirium at admission and on a daily basis after admission to the hospital. To aid delirium diagnosis, a review of medical and nursing notes for a full evaluation of delirium was performed by one of the investigators. Delirium was diagnosed according to the Diagnostic and Statistical Manual of Mental Disorders 5th edition. Delirium was defined as "early onset" if it was diagnosed within the first 24 hours after admission to the neurology unit due to acute ischemic stroke.			
		Dev	Val
3.49 Was the outcome determined appropriately?		Y	
3.50 Was a pre-specified or standard outcome definition used?		PY	
3.51 Were predictors excluded from the outcome definition?		PN	
3.52 Was the outcome defined and determined in a similar way for all participants?		PY	
3.53 Was the outcome determined without knowledge of predictor information?		N	
3.54 Was the time interval between predictor assessment and outcome determination appropriate?		NI	
Risk of bias introduced by the outcome or its determination	**RISK:** *(low/ high/ unclear)*	**H**	
Rationale of bias rating:			

It is unclear whether the outcome was assessed always after the predictors were determined; there are no details regarding the investigators experience/qualifications and the description of what criteria were used to diagnose delirium beyond use of CAM-ICU seems a bit vague, yet it seems that the investigator had knowledge of the predictors and they were used to aid judgement, which may have led to overestimation of associations between predictors and outcome.

B. Applicability

At what time point was the outcome determined:

At admission and on a daily basis afterwards up to 5 days.

If a composite outcome was used, describe the relative frequency/distribution of each contributing outcome:

N/A

Concern that the outcome, its definition, timing or determination do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	

Rationale of applicability rating:

The timing of outcome assessment is appropriate given that delirium can develop and recede shortly after admission; all other aspects of the outcome match the review question with CAM-ICU and DSM-5 criteria being appropriate for assessment of delirium.

DOMAIN 4: Analysis

Risk of Bias

Describe numbers of participants, number of candidate predictors, outcome events and events per candidate predictor:

1001 participants, seems like 40 candidate predictors?, 172 outcome events, EPV = 4.3

Describe how the model was developed (for example in regards to modelling technique (e.g. survival or logistic modelling), predictor selection, and risk group definition):

"The receiver operating characteristic (ROC) curve analysis was performed to determine the best cut-off value for predicting the clinical end points. Logistic regression analysis was performed, with additional correction for potentially interfering variables (age, sex, and body mass index (BMI)). On the basis of this regression, the parameters most closely related to the occurrence of delirium were sought. Based on a multifactorial model, information on the impact of each variable on delirium was obtained. Next, the contribution of each analyzed variable was calculated, and the formula was presented using elements with which the delirium score was calculated. p-value ≤ 0.05 was considered significant."

"Because patients with delirium had significantly higher values of NLR at admission, we decided to perform a ROC analysis to calculate the best cut-off value to predict delirium."

"Multivariable logistic regression analysis adjusted according to age, sex, BMI, comorbidities, and baseline neurology showed that leucocyte count (p < 0.001) and neutrophil count (p = 0.012) as well as mean NRL (p = 0.028), NRL (p < 0.001), and NLR at the predefined cut-off of >4.86 (p < 0.001) exhibited an association with post-stroke delirium after adjustment with baseline

characteristics. For NLR > 4.86 adjusted for age, sex, BMI, comorbidities, and baseline neurology, the odds ratio (OR) was 1.875 (95% CI 1.314–2.675, p = 0.001). Similar finding was noted regarding the CRP, with the cut-off >9.10, for which the OR adjusted for age, sex, BMI, comorbidities, and baseline neurology was even higher with NLR at 2.132 (95% CI 1.482–3.066, p < 0.001). The AUC value for NLR as well as its sensitivity and specificity were moderate, so we decided to find a combination of markers and clinical parameters that would help predict the occurrence of early-onset delirium in AIS. Using clinical and laboratory factors, we determined that an index composed of age, NIHSS score, neurological findings, leucocyte count, NLR, and CRP was better at predicting early-onset delirium after acute ischemic stroke than any of the factors alone. To increase the diagnostic value of the laboratory markers in the study group, clinical and laboratory variables most associated with delirium were determined based on logistic regression."

Describe whether and how the model was validated, either internally (e.g. bootstrapping, cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different setting, different type of participants):

"To verify the predictability value of the DELIAS score for the diagnosis of delirium, we performed an ROC analysis for delirium diagnosed up to the fifth day in the same group of patients."

Describe the performance measures of the model, e.g. (re)calibration, discrimination, (re)classification, net benefit, and whether they were adjusted for optimism:

Discrimination and classification measures.

Describe any participants who were excluded from the analysis:

Excluded patients with incomplete laboratory testing (6 patients), or no data regarding follow-up.

Describe missing data on predictors and outcomes as well as methods used for missing data:

Participants with missing data were excluded.

	Dev	Val
4.73 Were there a reasonable number of participants with the outcome?	N	
4.74 Were continuous and categorical predictors handled appropriately?	N	
4.75 Were all enrolled participants included in the analysis?	N	
4.76 Were participants with missing data handled appropriately?	N	
4.77 Was selection of predictors based on univariable analysis avoided?	NI	
4.78 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?	N	
4.79 Were relevant model performance measures evaluated appropriately?	N	
4.80 Were model overfitting and optimism in model performance accounted for?	N	
4.81 Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?	NI	

Risk of bias introduced by the analysis	RISK: *(low/ high/ unclear)*	H	
Rationale of bias rating: There were too few outcome events relative to candidate predictors; it is not clear how predictors were chosen or which were entered into the multivariable logistic model; participants with missing data were excluded, although this would most likely not cause a high ROB as this was only 19 from 1022; data-driven cut-offs for laboratory markers, only apparent validation, no information regarding calibration.			

Step 4: Overall assessment

Use the following tables to reach overall judgements about risk of bias and concerns regarding applicability of the prediction model evaluation (development and/or validation) across all assessed domains.
Complete for each evaluation of a distinct model.

Reaching an overall judgement about risk of bias of the prediction model evaluation	
Low risk of bias	If all domains were rated low risk of bias. If a prediction model was developed without any external validation, and it was rated as low risk of bias for all domains, consider downgrading to **high risk of bias**. Such a model can only be considered as low risk of bias, if the development was based on a very large data set and included some form of internal validation.
High risk of bias	If at least one domain is judged to be at **high risk of bias**.
Unclear risk of bias	If an unclear risk of bias was noted in at least one domain and it was low risk for all other domains.

Reaching an overall judgement about applicability of the prediction model evaluation	
Low concerns regarding applicability	If low concerns regarding applicability for all domains, the prediction model evaluation is judged to have **low concerns regarding applicability**.
High concerns regarding applicability	If high concerns regarding applicability for at least one domain, the prediction model evaluation is judged to have **high concerns regarding applicability**.
Unclear concerns regarding applicability	If unclear concerns (but no "high concern") regarding applicability for at least one domain, the prediction model evaluation is judged to have **unclear concerns regarding applicability** overall.

Overall judgement about risk of bias and applicability of the prediction model evaluation		
Overall judgement of risk of bias	**RISK:** *(low/ high/ unclear)*	**H**
Summary of sources of potential bias: Quite a lot of study aspects unclear, issues with outcome assessment involving knowledge of predictors, issues with analysis, assessment of performance and validation		
Overall judgement of applicability	**CONCERN:** *(low/ high/ unclear)*	**UN**
Summary of applicability concerns: Applicability cannot be fully assessed as there is insufficient information to determine whether all predictor data was collected prior to the outcome.		

Oldenbeuving et al., 2014

Step 2: Classify the type of prediction model evaluation

Use the following table to classify the evaluation as model development, model validation or model update, or combination. Different signalling questions apply for different types of prediction model evaluation. If the evaluation does not fit one of these classifications then PROBAST should not be used.

Classify the evaluation based on its aim			
Type of prediction study	**PROBAST boxes to complete**	**Tick as appropriate**	**Definition for type of prediction model study**
Development only	Development	✓	Prediction model development without external validation. These studies may include internal validation methods, such as bootstrapping and cross-validation techniques.
Development and validation	Development and validation	X	Prediction model development combined with external validation in other participants in the same article.
Validation only	Validation	X	External validation of existing (previously developed) model in other participants.

This table should be completed once for each publication being assessed and for each relevant outcome in your review.

Publication reference	Oldenbeuving, A. W. et al. An early prediction of delirium in the acute phase after stroke. 2014
Models of interest	Risk score based on demographic and clinical variables
Outcome of interest	Post-stroke delirium

Step 3: Assess risk of bias and applicability

PROBAST is structured as four key domains. Each domain is judged for risk of bias (low, high or unclear) and includes signalling questions to help make judgements. Signalling questions are rated as yes (Y), probably yes (PY), probably no (PN), no (N) or no information (NI). All signalling questions are phrased so that "yes" indicates absence of bias. Any signalling question rated as "no" or "probably no" flags the potential for bias; you will need to use your judgement to determine whether the domain should be rated as "high", "low" or "unclear" risk of bias. The guidance document contains further instructions and examples on rating signalling questions and risk of bias for each domain. The first three domains are also rated for concerns regarding applicability (low/ high/ unclear) to your review question defined above.
Complete all domains separately for each evaluation of a distinct model. Shaded boxes indicate where signalling questions do not apply and should not be answered.

DOMAIN 1: Participants		
A. Risk of Bias		
Describe the sources of data and criteria for participant selection: Prospective cohort; "During 1 year, 630 consecutive patients with stroke admitted to the stroke units of the St. Elisabeth and TweeSteden hospitals in Tilburg, the Netherlands, were investigated for the presence and risk factors of delirium. Criteria for stroke were neurologic deficit of sudden onset lasting longer than 24 hours. Patients with ischemic and hemorrhagic stroke were included. All patients were admitted to a stroke care unit and treated according to standard protocols according to international guidelines. Exclusion criteria. Patients with subarachnoid hemorrhage and TIA were excluded. Patients had to be older than 18 years. Of the 630 consecutive patients, 95 were excluded. Forty-four patients were excluded because they already died before the first screening or because death appeared imminent. One patient was younger than 18 years, 2 had severe mental retardation, 6 had a severe language barrier, 35 were transferred to another hospital because of capacity problems, and 7 patients refused informed consent. Eight patients were admitted twice in the same period; only the first admission was included in the analyses. Hence, 527 patients were included in the analysis."		
	Dev	Val
1.19 Were appropriate data sources used, e.g. cohort, RCT or nested case-control study data?	Y	
1.20 Were all inclusions and exclusions of participants appropriate?	Y	

Risk of bias introduced by selection of participants	**RISK:** *(low/ high/ unclear)*	**L**	
Rationale of bias rating: Overall, broad inclusion criteria mainly excluding participants with other conditions than that of interest or those for whom a cognitive assessment was inappropriate or not possible.			
B. Applicability			
Describe included participants, setting and dates: During 1 year, 630 consecutive patients with stroke admitted to the stroke units of the St. Elisabeth and TweeSteden hospitals in Tilburg, the Netherlands.			
Concern that the included participants and setting do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	
Rationale of applicability rating: Dates are unclear, however regardless the participants and setting match the review question.			

DOMAIN 2: Predictors		
A. Risk of Bias		
List and describe predictors included in the final model, e.g. definition and timing of assessment: Age, NIHSS , stroke subtype, infection "We collected the following baseline data: age, sex, medication at time of admission, alcohol use (defined as a mean intake of one or more units every day), and auditory and visual impairment. At admission, all patients underwent clinical examination and a noncontrast enhanced CT scan with 5-mm contiguous slices. Stroke subtype was classified with the Oxfordshire Community Stroke Project criteria. For the multivariable analysis, partial anterior circulation infarction (PACI) and total anterior circulation infarction (TACI) were grouped. The severity of the clinical deficits was scored according to the NIH Stroke Scale (NIHSS), both at admission and at the first screening for delirium. The NIHSS data of the first screening were used in the analyses." "Infection was scored at both screening dates using data from the medical records from the day of hospitalisation until the day of screening. We used the following data: pyrexia, high leucocytosis and/or raised ESR with a positive blood, sputum or urine culture and/or infiltrate on chest x-ray, or for which antibiotics were prescribed." "CT scans of 484 patients were analyzed for atrophy and white matter changes by 2 raters blinded for patient characteristics."		
	Dev	Val
2.28 Were predictors defined and assessed in a similar way for all participants?	N	
2.29 Were predictor assessments made without knowledge of outcome data?	PN	
2.30 Are all predictors available at the time the model is intended to be used?	N	

Risk of bias introduced by predictors or their assessment	**RISK:** *(low/ high/ unclear)*	**H**	
Rationale of bias rating: Different variables were used for determining infection, although this seems appropriate given potential differences in type and symptoms of infection. Instead of using NIHSS from admission, a score that is determined on the same occasion as delirium was included, which could have led to overestimation of the association between this predictor and the outcome. Also, infection could have been recorded after a diagnosis of delirium was made at the first screening.			
B. Applicability			
Concern that the definition, assessment or timing of predictors in the model do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**H**	
Rationale of applicability rating: Issue due to overlap of predictor and outcome assessment, as stated above.			

DOMAIN 3: Outcome			
A. Risk of Bias			
Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination: "Every patient was screened for delirium between days 2 and 4 after admission, and a second time between days 5 and 7. If the patient was discharged before the second delirium screening, only the first screening was performed. Delirium was assessed with the Confusion Assessment Method (CAM). If the CAM was positive, delirium was diagnosed and the severity of delirium was assessed daily with the Delirium Rating Scale (DRS)."			
		Dev	Val
3.55 Was the outcome determined appropriately?		N	
3.56 Was a pre-specified or standard outcome definition used?		Y	
3.57 Were predictors excluded from the outcome definition?		Y	
3.58 Was the outcome defined and determined in a similar way for all participants?		Y	
3.59 Was the outcome determined without knowledge of predictor information?		N	
3.60 Was the time interval between predictor assessment and outcome determination appropriate?		N	
Risk of bias introduced by the outcome or its determination	**RISK:** *(low/ high/ unclear)*	**H**	
Rationale of bias rating: The outcome measure itself was appropriate, however the timing of assessment could have led to missing occurrence of delirium, particularly as the authors reported that it could last just 1 day and some patients were not screened till day 4 after the admission, it is also unclear whether the interval between the 1st and 2nd screenings was random or predetermined, e.g. could one patient be screened at days 4 and 5? NIHSS was assessed at the same time as the outcome and infection could have been recorded even after that,			

which indicates that predictors were known to the person determining the outcome; there is also no mention of blinding to any other of the predictors collected at admission, this introduces a high ROB as assessment of CAM involves some subjectivity.

B. Applicability

At what time point was the outcome determined:

Between days 2 and 4 after admission, and a second time between days 5 and 7.

If a composite outcome was used, describe the relative frequency/distribution of each contributing outcome:

N/A

Concern that the outcome, its definition, timing or determination do not match the review question	**CONCERN:** *(low/ high/ unclear)*	**L**	

Rationale of applicability rating:

Despite issues described above, all aspects of the outcome match the review question.

DOMAIN 4: Analysis

Risk of Bias

Describe numbers of participants, number of candidate predictors, outcome events and events per candidate predictor:

527 participants, 18 candidate predictors, 62 outcome events, EPV = 3.4

Describe how the model was developed (for example in regards to modelling technique (e.g. survival or logistic modelling), predictor selection, and risk group definition):

From preceding study:
A multivariable logistic regression analysis was performed with delirium as dependent variable. Initially, a full model was analyzed including all variables with a p value of 0.20 in the univariable analysis. For the multivariable analyses, a backward elimination procedure was used to define the final independent risk factors. Variables were eliminated from the model if the p value was 0.10.
From current study:
Age was an independent risk factor if brain atrophy was left out of the model. "For the current study, our aim was to develop a risk score that is available on the day of admission and that can be easily obtained. Therefore, we used age instead of brain atrophy. By means of the ß coefficients from the logistic regression model, we allocated a score to each risk factor. Three models were tested: 1 with all variables found significant in the multivariable analysis in the previous study with age instead of atrophy, 2 including only variables easily available for clinicians, 3 further simplified with only age and NIHSS"

Describe whether and how the model was validated, either internally (e.g. bootstrapping, cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different setting, different type of participants):

Apparent and temporal validation

Describe the performance measures of the model, e.g. (re)calibration, discrimination, (re)classification, net benefit, and whether they were adjusted for optimism:

Discrimination and classification measures

Describe any participants who were excluded from the analysis:

Forty-four patients were excluded because they already died before the first screening or because death appeared imminent, 35 were transferred to another hospital because of capacity problems

Describe missing data on predictors and outcomes as well as methods used for missing data:

No mention of missing data

		Dev	Val
4.82 Were there a reasonable number of participants with the outcome?		N	
4.83 Were continuous and categorical predictors handled appropriately?		PN	
4.84 Were all enrolled participants included in the analysis?		N	
4.85 Were participants with missing data handled appropriately?		N	
4.86 Was selection of predictors based on univariable analysis avoided?		N	
4.87 Were complexities in the data (e.g. censoring, competing risks, sampling of controls) accounted for appropriately?		N	
4.88 Were relevant model performance measures evaluated appropriately?		N	
4.89 Were model overfitting and optimism in model performance accounted for?		N	
4.90 Do predictors and their assigned weights in the final model correspond to the results from multivariable analysis?		NI	
Risk of bias introduced by the analysis	**RISK:** *(low/ high/ unclear)*	**H**	

Rationale of bias rating:

Too few participants relative to number of candidate predictors, not clear how NIHSS categories were created, as they do not align with typically applied cut-offs, participants who died or were lost to follow-up were excluded from the study, predictors were initially selected based on univariable analysis, no assessment of calibration, apparent validation with no correction for optimism, beta coefficients used to create the final score were not presented in either publication

Step 4: Overall assessment

Use the following tables to reach overall judgements about risk of bias and concerns regarding applicability of the prediction model evaluation (development and/or validation) across all assessed domains.
Complete for each evaluation of a distinct model.

Reaching an overall judgement about risk of bias of the prediction model evaluation	
Low risk of bias	If all domains were rated low risk of bias. If a prediction model was developed without any external validation, and it was rated as low risk of bias for all domains, consider downgrading to **high risk of bias**. Such a model can only be considered as low risk of bias, if the development was based on a very large data set and included some form of internal validation.
High risk of bias	If at least one domain is judged to be at **high risk of bias**.
Unclear risk of bias	If an unclear risk of bias was noted in at least one domain and it was low risk for all other domains.

Reaching an overall judgement about applicability of the prediction model evaluation	
Low concerns regarding applicability	If low concerns regarding applicability for all domains, the prediction model evaluation is judged to have **low concerns regarding applicability**.
High concerns regarding applicability	If high concerns regarding applicability for at least one domain, the prediction model evaluation is judged to have **high concerns regarding applicability**.
Unclear concerns regarding applicability	If unclear concerns (but no "high concern") regarding applicability for at least one domain, the prediction model evaluation is judged to have **unclear concerns regarding applicability** overall.

Overall judgement about risk of bias and applicability of the prediction model evaluation		
Overall judgement of risk of bias	**RISK:** *(low/ high/ unclear)*	**H**
Summary of sources of potential bias: Issues in terms of overlap in assessment of predictors and outcome with inherent lack of blinding, timing of outcome assessment may have led to missing occurrences of delirium, multiple issues regarding model development and assessment of performance		
Overall judgement of applicability	**CONCERN:** *(low/ high/ unclear)*	**H**
Summary of applicability concerns: Review question specifies that predictors should be known prior to the occurrence of the outcome, which here was not the case.		

Appendix 4: Chapter 4, associations between performance on cognitive tasks and measures of physical activity and sedentary behaviour in the UK Biobank

Supplemental Table 1 Associations between log reaction time and daily duration of types of physical activities and sedentary behaviours.

	Log reaction time models							
	Unadjusted		Partially adjusted		Fully adjusted		Complete	
Variable	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)
Physical activities (mins/day)								
Walking[a]	−0.011* (−0.019, −0.003)	−0.061* (−0.103, −0.019)	−0.011* (−0.019, −0.003)	−0.059* (−0.100, −0.017)	−0.005 (−0.015, 0.005)	−0.027 (−0.080, 0.027)	−0.005 (−0.015, 0.005)	−0.029 (−0.082, 0.024)
Moderate activity[a]	0.001 (−0.005, 0.007)	0.005 (−0.043, 0.054)	−0.001 (−0.007, 0.004)	−0.012 (−0.060, 0.036)	<0.001 (−0.007, 0.007)	0.001 (−0.059, 0.060)	<0.001 (−0.007, 0.008)	0.003 (−0.057, 0.063)
Vigorous activity[a]	−0.006* (−0.012, −0.001)	−0.059* (−0.110, −0.008)	−0.003 (−0.008, 0.002)	−0.030 (−0.080, 0.021)	<0.001 (−0.006, 0.007)	0.002 (−0.060, 0.065)	<0.001 (−0.006, 0.007)	0.002 (−0.061, 0.064)
Sample	5,739		5,675		3,574		3,535	
Sedentary behaviours (hrs/day)								
Driving[a]	−0.032* (−0.047, −0.017)	−0.088* (−0.128, −0.048)	−0.020* (−0.035, −0.005)	−0.054* (−0.095, −0.013)	−0.007 (−0.026, 0.011)	−0.021 (−0.071, 0.030)	−0.012 (−0.033, 0.010)	−0.032 (−0.090, 0.027)
Computer use[a]	−0.035* (−0.048, −0.023)	−0.104* (−0.140, −0.067)	−0.028* (−0.040, −0.015)	−0.081* (−0.118, −0.044)	−0.023* (−0.039, −0.008)	−0.069* (−0.115, −0.023)	−0.024* (−0.042, −0.006)	−0.070* (−0.123, −0.017)
Watching TV	0.008* (0.005, 0.012)	0.075* (0.045, 0.105)	0.006* (0.002, 0.009)	0.052* (0.021, 0.082)	0.003 (−0.002, 0.008)	0.027 (−0.017, 0.070)	0.003 (−0.002, 0.009)	0.028 (−0.022, 0.078)
Sample	7,771		7,673		4,564		3,535	

*Significant at $p < 0.003$; [a]natural log of reported duration

CI indicates confidence interval; hrs, hours; mins, minutes; std., standardised; unstd., unstandardised.

Supplemental Table 2 Associations between log reaction time and total daily physically active and sedentary times.

	Log reaction time models							
	Unadjusted		Partially adjusted		Fully adjusted		Complete	
Variable	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)
Total active time[a] (mins/day)	−0.012* (−0.018, −0.005)	−0.064* (−0.099, −0.029)	−0.011* (−0.018, −0.005)	−0.061* (−0.095, −0.026)	−0.002 (−0.011, 0.007)	−0.010 (−0.059, 0.038)	−0.002 (−0.011, 0.007)	−0.011 (−0.060, 0.037)
Sample	5,739		5,675		3,574		3,535	
Total sedentary time (hrs/day)	−0.001 (−0.003, 0.002)	−0.007 (−0.042, 0.028)	−0.001 (−0.003, 0.002)	−0.007 (−0.043, 0.028)	−0.002 (−0.005, 0.002)	−0.021 (−0.069, 0.027)	−0.002 (−0.006, 0.002)	−0.025 (−0.080, 0.030)
Sample	7,770		7,672		4,564		3,535	

*Significant at $p < 0.003$; [a]natural log of reported duration

CI indicates confidence interval; hrs, hours; mins, minutes; std., standardised; unstd., unstandardised.

Supplemental Table 3 Associations between verbal-numerical reasoning task scores and daily duration of types of physical activities and sedentary behaviours.

	Verbal-numerical reasoning models							
	Unadjusted		Partially adjusted		Fully adjusted		Complete	
Variable	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)
Physical activities (mins/day)								
Walking[a]	−0.023 (−0.164, 0.117)	−0.011 (−0.078, 0.056)	−0.023 (−0.158, 0.113)	−0.011 (−0.076, 0.054)	−0.075 (−0.224, 0.074)	−0.036 (−0.107, 0.036)	−0.053 (−0.203, 0.096)	−0.025 (−0.097, 0.046)
Moderate activity[a]	0.015 (−0.089, 0.120)	0.011 (−0.065, 0.087)	0.036 (−0.066, 0.138)	0.026 (−0.048, 0.100)	−0.013 (−0.124, 0.097)	−0.010 (−0.090, 0.071)	−0.021 (−0.131, 0.090)	−0.015 (−0.095, 0.065)
Vigorous activity[a]	0.045 (−0.045, 0.135)	0.038 (−0.038, 0.114)	−0.014 (−0.102, 0.074)	−0.012 (−0.086, 0.063)	−0.033 (−0.128, 0.063)	−0.028 (−0.109, 0.053)	−0.030 (−0.125, 0.066)	−0.025 (−0.106, 0.056)
Sample	1,927		1,906		1,583		1,571	
Sedentary behaviours (hrs/day)								
Driving[a]	−0.142 (−0.388, 0.104)	−0.035 (−0.094, 0.025)	−0.162 (−0.409, 0.085)	−0.039 (−0.099, 0.021)	−0.202 (−0.478, 0.075)	−0.049 (−0.116, 0.018)	−0.267 (−0.585, 0.052)	−0.065 (−0.142, 0.013)
Computer use[a]	0.546* (0.342, 0.749)	0.142* (0.089, 0.195)	0.396* (0.192, 0.599)	0.103* (0.050, 0.156)	0.336* (0.108, 0.564)	0.088* (0.028, 0.147)	0.275* (0.018, 0.533)	0.072* (0.005, 0.139)
Watching TV	−0.191* (−0.248, −0.133)	−0.151* (−0.197, −0.106)	−0.133* (−0.191, −0.075)	−0.106* (−0.152, −0.060)	−0.100* (−0.169, −0.032)	−0.080* (−0.134, −0.025)	−0.116* (−0.194, −0.038)	−0.092* (−0.154, −0.030)
Sample	2,530		2,500		1,975		1,571	

*Significant at $p < 0.003$; [a]natural log of reported duration

CI indicates confidence interval; hrs, hours; mins, minutes; std., standardised; unstd., unstandardised.

Supplemental Table 4 Associations between verbal-numerical reasoning task scores and total daily physically active and sedentary times.

	Verbal-numerical reasoning models							
	Unadjusted		Partially adjusted		Fully adjusted		Complete	
Variable	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)
Total active time[a] (mins/day)	0.049 (−0.069, 0.166)	0.023 (−0.033, 0.080)	0.018 (−0.095, 0.132)	0.009 (−0.046, 0.063)	−0.103 (−0.238, 0.032)	−0.050 (−0.115, 0.015)	−0.109 (−0.244, 0.026)	−0.053 (−0.118, 0.013)
Sample	1,927		1,906		1,583		1,571	
Total sedentary time (hrs/day)	−0.068* (−0.113, −0.022)	−0.081* (−0.134, −0.027)	−0.048* (−0.092, −0.003)	−0.057* (−0.109, −0.004)	−0.038 (−0.089, 0.014)	−0.045 (−0.106, 0.016)	−0.056 (−0.115, 0.003)	−0.066 (−0.136, 0.004)
Sample	2,530		2,500		1,975		1,571	

*Significant at $p < 0.003$; [a]natural log of reported duration

CI indicates confidence interval; hrs, hours; mins, minutes; std., standardised; unstd., unstandardised.

Supplemental Table 5 Associations between log errors in the visual memory task and daily duration of types of physical activities and sedentary behaviours.

	Visual memory models (log of errors)							
	Unadjusted		Partially adjusted		Fully adjusted		Complete	
Variable	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)
Physical activities (mins/day)								
Walking[a]	−0.004 (−0.029, 0.020)	−0.007 (−0.046, 0.032)	−0.004 (−0.029, 0.021)	−0.007 (−0.045, 0.032)	−0.006 (−0.038, 0.026)	−0.009 (−0.059, 0.041)	−0.007 (−0.040, 0.025)	−0.012 (−0.062, 0.039)
Moderate activity[a]	0.019* (0.000, 0.037)	0.044* (0.000, 0.089)	0.015 (−0.004, 0.034)	0.035 (−0.010, 0.079)	0.018 (−0.006, 0.042)	0.042 (−0.014, 0.098)	0.019 (−0.004, 0.043)	0.046 (−0.010, 0.102)
Vigorous activity[a]	0.005 (−0.012, 0.022)	0.013 (−0.034, 0.060)	0.009 (−0.008, 0.026)	0.024 (−0.023, 0.072)	−0.002 (−0.023, 0.020)	−0.005 (−0.064, 0.054)	−0.003 (−0.025, 0.018)	−0.009 (−0.068, 0.050)
Sample	5,845		5,782		3,626		3,584	
Sedentary behaviours (hrs/day)								
Driving[a]	0.005 (−0.042, 0.053)	0.004 (−0.033, 0.042)	0.015 (−0.034, 0.063)	0.011 (−0.027, 0.050)	0.015 (−0.046, 0.077)	0.012 (−0.036, 0.061)	−0.005 (−0.075, 0.066)	−0.004 (−0.060, 0.052)
Computer use[a]	−0.013 (−0.053, 0.027)	−0.011 (−0.045, 0.023)	−0.006 (−0.046, 0.035)	−0.005 (−0.039, 0.030)	−0.018 (−0.070, 0.034)	−0.015 (−0.060, 0.029)	−0.014 (−0.073, 0.045)	−0.012 (−0.062, 0.038)
Watching TV	−0.008 (−0.019, 0.003)	−0.021 (−0.049, 0.006)	−0.011 (−0.022, <0.001)	−0.028 (−0.057, <0.001)	−0.009 (−0.025, 0.007)	−0.022 (−0.064, 0.019)	−0.002 (−0.020, 0.017)	−0.004 (−0.051, 0.043)
Sample	7,927		7,825		4,639		3,584	

*Significant at $p < 0.003$; [a]natural log of reported duration

CI indicates confidence interval; hrs, hours; mins, minutes; std., standardised; unstd., unstandardised.

Supplemental Table 6 Associations between log errors in the visual memory task and total daily physically active and sedentary times.

	Visual memory models (log of errors)							
	Unadjusted		Partially adjusted		Fully adjusted		Complete	
Variable	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)
Total active time[a] (mins/day)	0.020 (−0.001, 0.040)	0.031 (−0.001, 0.063)	0.018 (−0.002, 0.039)	0.029 (−0.003, 0.061)	0.011 (−0.018, 0.040)	0.017 (−0.029, 0.062)	0.011 (−0.018, 0.040)	0.017 (−0.029, 0.062)
Sample	5,845		5,782		3,626		3,584	
Total sedentary time (hrs/day)	−0.007 (−0.016, 0.001)	−0.029 (−0.061, 0.004)	−0.007 (−0.016, 0.001)	−0.029 (−0.062, 0.004)	−0.006 (−0.018, 0.006)	−0.024 (−0.070, 0.021)	−0.004 (−0.017, 0.010)	−0.015 (−0.067, 0.038)
Sample	7,926		7,824		4,639		3,584	

*Significant at $p < 0.003$; [a]natural log of reported duration

CI indicates confidence interval; hrs, hours; mins, minutes; std., standardised; unstd., unstandardised.

Supplemental Table 7 Associations between a correct response on the prospective memory task and daily duration of types of physical activities and sedentary behaviours.

	Prospective memory models							
	Unadjusted		Partially adjusted		Fully adjusted		Complete	
Variable	Unstd. beta (99.7% CI)	Odds ratio (99.7% CI)	Unstd. beta (99.7% CI)	Odds ratio (99.7% CI)	Unstd. beta (99.7% CI)	Odds ratio (99.7% CI)	Unstd. beta (99.7% CI)	Odds ratio (99.7% CI)
Physical activities (mins/day)								
Walking[a]	0.008 (−0.130, 0.146)	1.008 (0.878, 1.157)	0.003 (−0.137, 0.143)	1.003 (0.872, 1.154)	−0.043 (−0.206, 0.120)	0.958 (0.814, 1.127)	−0.037 (−0.203, 0.128)	0.963 (0.817, 1.136)
Moderate activity[a]	0.029 (−0.075, 0.132)	1.029 (0.928, 1.141)	0.047 (−0.059, 0.153)	1.048 (0.943, 1.166)	0.014 (−0.108, 0.136)	1.014 (0.897, 1.146)	0.006 (−0.118, 0.129)	1.006 (0.889, 1.138)
Vigorous activity[a]	−0.003 (−0.093, 0.087)	0.997 (0.911, 1.090)	−0.031 (−0.125, 0.062)	0.969 (0.883, 1.064)	−0.048 (−0.154, 0.059)	0.953 (0.857, 1.061)	−0.041 (−0.149, 0.066)	0.959 (0.862, 1.068)
Sample	2,037		2,014		1,660		1,644	
Sedentary behaviours (hrs/day)								
Driving[a]	−0.137 (−0.385, 0.111)	0.872 (0.680, 1.117)	−0.178 (−0.436, 0.080)	0.837 (0.647, 1.084)	−0.104 (−0.405, 0.198)	0.901 (0.667, 1.218)	−0.043 (−0.395, 0.309)	0.958 (0.674, 1.362)
Computer use[a]	0.275* (0.064, 0.486)	1.316* (1.066, 1.625)	0.224* (0.005, 0.443)	1.251* (1.005, 1.557)	0.234 (−0.022, 0.490)	1.263 (0.978, 1.632)	0.171 (−0.118, 0.460)	1.187 (0.889, 1.584)
Watching TV	−0.097* (−0.154, −0.039)	0.908* (0.857, 0.962)	−0.063* (−0.123, −0.004)	0.939* (0.884, 0.996)	−0.026 (−0.101, 0.048)	0.974 (0.904, 1.049)	−0.045 (−0.130, 0.040)	0.956 (0.876, 1.039)
Sample	2,709		2,669		2,081		1,644	

*Significant at $p < 0.003$; [a]natural log of reported duration

CI indicates confidence interval; hrs, hours; mins, minutes; std., standardised; unstd., unstandardised.

Supplemental Table 8 Associations between a correct response on the prospective memory task and total daily physically active and sedentary times.

	Prospective memory models							
	Unadjusted		Partially adjusted		Fully adjusted		Complete	
Variable	Unstd. beta (99.7% CI)	Odds ratio (99.7% CI)	Unstd. beta (99.7% CI)	Odds ratio (99.7% CI)	Unstd. beta (99.7% CI)	Odds ratio (99.7% CI)	Unstd. beta (99.7% CI)	Odds ratio (99.7% CI)
Total active time[a] (mins/day)	0.040 (−0.073, 0.152)	1.041 (0.930, 1.165)	0.032 (−0.083, 0.146)	1.032 (0.921, 1.157)	−0.049 (−0.195, 0.098)	0.952 (0.823, 1.102)	−0.052 (−0.199, 0.095)	0.949 (0.819, 1.099)
Sample	2,037		2,014		1,660		1,644	
Total sedentary time (hrs/day)	−0.037 (−0.081, 0.006)	0.964 (0.923, 1.007)	−0.025 (−0.070, 0.020)	0.975 (0.932, 1.021)	0.003 (−0.052, 0.059)	1.003 (0.949, 1.060)	−0.008 (−0.072, 0.056)	0.992 (0.931, 1.058)
Sample	2,709		2,669		2,081		1,644	

*Significant at $p < 0.003$; [a]natural log of reported duration

CI indicates confidence interval; hrs, hours; mins, minutes; std., standardised; unstd., unstandardised.

Appendix 5: Chapter 4, associations between performance on cognitive tasks and proxies of social engagement in the UK Biobank

Supplemental Table 9 Associations between log reaction time and proxies of social engagement.

	Log reaction time models							
	Unadjusted		Partially adjusted		Fully adjusted		Complete	
Variable	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)
Family/friend visits (never[a])								
Monthly	−0.052* (−0.095, −0.009)	−0.272* (−0.495, −0.049)	−0.042 (−0.084, 0.001)	−0.219 (−0.440, 0.003)	−0.063* (−0.119, −0.007)	−0.330* (−0.622, −0.038)	−0.066 (−0.159, 0.026)	−0.346 (−0.828, 0.135)
Weekly	−0.046* (−0.086, −0.006)	−0.242* (−0.451, −0.032)	−0.043* (−0.083, −0.003)	−0.226* (−0.435, −0.018)	−0.051 (−0.104, 0.002)	−0.266 (−0.545, 0.012)	−0.049 (−0.138, 0.040)	−0.256 (−0.719, 0.207)
Daily	−0.028 (−0.071, 0.015)	−0.145 (−0.371, 0.080)	−0.032 (−0.075, 0.011)	−0.165 (−0.389, 0.060)	−0.047 (−0.104, 0.010)	−0.244 (−0.542, 0.054)	−0.045 (−0.140, 0.050)	−0.235 (−0.730, 0.261)
Sample	8,052		7,934		4,689		1,954	
Family relationship satisfaction (not satisfied[a])								
Satisfied	−0.043 (−0.087, 0.001)	−0.226 (−0.455, 0.003)	−0.051* (−0.095, −0.007)	−0.266* (−0.494, −0.038)	−0.042 (−0.095, 0.011)	−0.219 (−0.496, 0.059)	−0.038 (−0.096, 0.020)	−0.197 (−0.499, 0.105)
Sample	2,744		2,696		2,093		1,954	
Friendship satisfaction (not satisfied[a])								
Satisfied	−0.008 (−0.069, 0.053)	−0.042 (−0.363, 0.279)	−0.021 (−0.081, 0.040)	−0.108 (−0.424, 0.208)	−0.034 (−0.112, 0.044)	−0.180 (−0.588, 0.229)	−0.024 (−0.109, 0.060)	−0.128 (−0.568, 0.313)
Sample	2,731		2,685		2,085		1,954	
Loneliness (not lonely[a])								
Lonely	0.024* (0.008, 0.041)	0.127* (0.040, 0.214)	0.020* (0.003, 0.037)	0.105* (0.017, 0.194)	0.007 (−0.017, 0.031)	0.038 (−0.089, 0.164)	−0.013 (−0.051, 0.024)	−0.070 (−0.267, 0.126)
Sample	8,033		7,847		4,637		1,954	

Supplemental Table 9 Associations between log reaction time and proxies of social engagement. ***Continued***

Variable	Unadjusted		Partially adjusted		Fully adjusted		Complete	
	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)
Confiding in someone (never[a])								
Monthly	0.020 (−0.007, 0.047)	0.105 (−0.034, 0.244)	0.022 (−0.005, 0.048)	0.114 (−0.024, 0.252)	0.027 (−0.007, 0.061)	0.140 (−0.035, 0.316)	−0.001 (−0.054, 0.053)	−0.003 (−0.282, 0.276)
Weekly	0.002 (−0.021, 0.025)	0.010 (−0.112, 0.133)	0.004 (−0.019, 0.027)	0.021 (−0.102, 0.143)	0.010 (−0.020, 0.039)	0.050 (−0.105, 0.205)	0.009 (−0.039, 0.057)	0.045 (−0.206, 0.296)
Daily	−0.003 (−0.022, 0.016)	−0.017 (−0.117, 0.084)	0.001 (−0.019, 0.020)	0.003 (−0.098, 0.103)	0.007 (−0.017, 0.031)	0.034 (−0.091, 0.160)	0.001 (−0.041, 0.044)	0.008 (−0.213, 0.228)
Sample	7,852		7,674		4,543		1,954	
Social activities (none[a])								
Sports	−0.031* (−0.058, −0.003)	-0.160* (−0.304, −0.016)	−0.018 (−0.045, 0.009)	−0.094 (−0.237, 0.050)	0.002 (−0.031, 0.035)	0.011 (−0.161, 0.183)	0.005 (−0.046, 0.056)	0.028 (−0.239, 0.295)
Pub/social club	0.008 (−0.013, 0.029)	0.043 (−0.068, 0.154)	0.005 (−0.016, 0.026)	0.027 (−0.085, 0.138)	0.005 (−0.022, 0.032)	0.025 (−0.117, 0.166)	0.002 (−0.043, 0.047)	0.011 (−0.224, 0.246)
Religious group	0.055* (0.026, 0.085)	0.289* (0.136, 0.442)	0.048* (0.019 0.077)	0.249* (0.097, 0.402)	0.047* (0.011, 0.084)	0.248* (0.056, 0.440)	0.040 (−0.017, 0.097)	0.209 (−0.091, 0.509)
Adult education	0.021 (−0.034, 0.075)	0.109 (−0.175, 0.394)	0.031 (−0.023, 0.085)	0.162 (−0.121, 0.445)	−0.003 (−0.075, 0.069)	−0.015 (−0.392, 0.362)	0.009 (−0.111, 0.129)	0.046 (−0.582, 0.675)
Other	0.013 (−0.013, 0.038)	0.065 (−0.070, 0.201)	0.010 (−0.016, 0.035)	0.050 (−0.084, 0.185)	−0.005 (−0.037, 0.028)	−0.024 (−0.195, 0.146)	0.006 (−0.046, 0.059)	0.032 (−0.242, 0.307)
Multiple	−0.002 (−0.022, 0.017)	−0.013 (−0.115, 0.090)	−0.001 (−0.021, 0.019)	−0.005 (−0.108, 0.097)	0.014 (−0.011, 0.038)	0.072 (−0.057, 0.201)	0.019 (−0.021, 0.058)	0.098 (−0.107, 0.303)
Sample	8,152		7,968		4,698		1,954	

*Significant at $p < 0.003$; [a]reference category

CI indicates confidence interval; std., standardised; unstd., unstandardised.

Supplemental Table 10 Associations between verbal-numerical reasoning task scores and proxies of social engagement.

	Verbal-numerical reasoning models							
	Unadjusted		Partially adjusted		Fully adjusted		Complete	
Variable	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)
Family/friend visits (never[a])								
Monthly	0.373 (−0.352, 1.099)	0.173 (−0.163, 0.509)	0.311 (−0.392, 1.014)	0.144 (−0.181, 0.470)	0.395 (−0.402, 1.191)	0.183 (−0.186, 0.551)	0.391 (−0.472, 1.254)	0.181 (−0.218, 0.581)
Weekly	0.332 (−0.351, 1.014)	0.154 (−0.163, 0.470)	0.250 (−0.413, 0.913)	0.116 (−0.191, 0.423)	0.317 (−0.438, 1.072)	0.147 (−0.203, 0.497)	0.271 (−0.559, 1.101)	0.125 (−0.259, 0.510)
Daily	−0.025 (−0.763, 0.712)	−0.012 (−0.353, 0.330)	−0.005 (−0.720, 0.711)	−0.002 (−0.333, 0.329)	0.056 (−0.759, 0.871)	0.026 (−0.352, 0.403)	0.029 (−0.861, 0.918)	0.013 (−0.399, 0.425)
Sample	2,611		2,572		2,018		1,873	
Family relationship satisfaction (not satisfied[a])								
Satisfied	0.282 (−0.155, 0.719)	0.130 (−0.072, 0.333)	0.241 (−0.180, 0.663)	0.112 (−0.083, 0.307)	0.125 (−0.391, 0.641)	0.058 (−0.181, 0.297)	−0.068 (−0.620, 0.485)	−0.031 (−0.287, 0.225)
Sample	2,580		2,541		2,000		1,873	
Friendship satisfaction (not satisfied[a])								
Satisfied	0.380 (−0.231, 0.991)	0.176 (−0.107, 0.459)	0.384 (−0.201, 0.969)	0.178 (−0.093, 0.449)	0.230 (−0.528, 0.988)	0.107 (−0.244, 0.458)	0.257 (−0.550, 1.064)	0.119 (−0.255, 0.493)
Sample	2,573		2,534		1,995		1,873	
Loneliness (not lonely[a])								
Lonely	−0.654* (−0.938, −0.370)	−0.303* (−0.434, −0.171)	−0.417* (−0.697, −0.136)	−0.193* (−0.323, −0.063)	−0.412* (−0.750, −0.074)	−0.191* (−0.347, −0.034)	−0.345 (−0.704, 0.013)	−0.160 (−0.326, 0.006)
Sample	2,582		2,543		1,997		1,873	

Supplemental Table 10 Associations between verbal-numerical reasoning task scores and proxies of social engagement. ***Continued***

Variable	Unadjusted		Partially adjusted		Fully adjusted		Complete	
	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)
Confiding in someone (never[a])								
Monthly	−0.089 (−0.537, 0.358)	−0.041 (−0.249, 0.166)	−0.228 (−0.658, 0.206)	−0.105 (−0.305, 0.095)	−0.227 (−0.713, 0.259)	−0.105 (−0.330, 0.120)	−0.279 (−0.780, 0.222)	−0.129 (−0.361, 0.103)
Weekly	0.022 (−0.376, 0.420)	0.010 (−0.174, 0.194)	−0.023 (−0.409, 0.363)	−0.011 (−0.189, 0.168)	−0.038 (−0.473, 0.397)	−0.018 (−0.219, 0.184)	−0.151 (−0.602, 0.300)	−0.070 (−0.279, 0.139)
Daily	0.087 (−0.256, 0.431)	0.040 (−0.119, 0.199)	0.017 (−0.315, 0.350)	0.008 (−0.146, 0.162)	0.052 (−0.322, 0.427)	0.024 (−0.149, 0.198)	−0.038 (−0.434, 0.358)	−0.017 (−0.201, 0.166)
Sample	2,525		2,488		1,963		1,873	
Social activities (none[a])								
Sports	0.379 (−0.071, 0.829)	0.175 (−0.033, 0.384)	0.123 (−0.311, 0.557)	0.057 (−0.144, 0.258)	0.046 (−0.423, 0.516)	0.022 (−0.196, 0.239)	0.010 (−0.472, 0.491)	0.005 (−0.218, 0.228)
Pub/social club	−0.199 (−0.569, 0.170)	−0.092 (−0.263, 0.079)	−0.119 (−0.478, 0.240)	−0.055 (−0.222, 0.111)	−0.284 (−0.700, 0.131)	−0.132 (−0.324, 0.061)	−0.338 (−0.765, 0.089)	−0.156 (−0.354, 0.041)
Religious group	−0.132 (−0.628, 0.363)	−0.061 (−0.291, 0.168)	−0.202 (−0.680, 0.277)	−0.093 (−0.315, 0.128)	−0.144 (−0.681, 0.392)	−0.067 (−0.315, 0.182)	−0.064 (−0.618, 0.489)	−0.030 (−0.286, 0.227)
Adult education	0.052 (−0.941, 1.045)	0.024 (−0.436, 0.484)	−0.035 (−0.981, 0.911)	−0.016 (−0.454, 0.422)	0.120 (−0.919, 1.160)	0.056 (−0.426, 0.537)	0.208 (−0.907, 1.323)	0.096 (−0.420, 0.612)
Other	0.238 (−0.212, 0.688)	0.110 (−0.098, 0.319)	0.225 (−0.209, 0.659)	0.104 (−0.097, 0.305)	0.218 (−0.265, 0.700)	0.101 (−0.123, 0.324)	0.268 (−0.225, 0.760)	0.124 (−0.104, 0.352)
Multiple	0.530* (0.200, 0.861)	0.246* (0.093, 0.398)	0.277 (−0.043, 0.597)	0.128 (−0.020, 0.277)	0.058 (−0.301, 0.418)	0.027 (−0.139, 0.193)	0.055 (−0.315, 0.425)	0.025 (−0.146, 0.197)
Sample	2,614		2,574		2,020		1,873	

*Significant at $p < 0.003$; [a]reference category

CI indicates confidence interval; std., standardised; unstd., unstandardised.

Supplemental Table 11 Associations between log errors in the visual memory task and proxies of social engagement.

	Visual memory models (log of errors)							
	Unadjusted		Partially adjusted		Fully adjusted		Complete	
Variable	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)
Family/friend visits (never[a])								
Monthly	−0.016 (−0.151, 0.119)	−0.024 (−0.229, 0.180)	−0.015 (−0.151, 0.122)	−0.022 (−0.229, 0.184)	−0.059 (−0.243, 0.125)	−0.089 (−0.367, 0.188)	−0.011 (−0.305, 0.283)	−0.017 (−0.462, 0.427)
Weekly	−0.014 (−0.141, 0.113)	−0.021 (−213, 0.171)	−0.022 (−0.150, 0.107)	−0.033 (−0.227, 0.162)	−0.091 (−0.266, 0.084)	−0.137 (−0.402, 0.127)	−0.088 (−0.370, 0.195)	−0.132 (−0.560, 0.295)
Daily	0.001 (−0.136, 0.137)	0.001 (−0.205, 0.208)	−0.005 (−0.143, 0.133)	−0.007 (−0.216, 0.202)	−0.074 (−0.261, 0.113)	−0.112 (−0.395, 0.171)	−0.112 (−0.414, 0.190)	−0.170 (−0.627, 0.288)
Sample	8,232		8,108		4,769		1,991	
Family relationship satisfaction (not satisfied[a])								
Satisfied	−0.072 (−0.213, 0.068)	−0.110 (−0.322, 0.103)	−0.102 (−0.244, 0.039)	−0.155 (−0.369, 0.060)	−0.110 (−0.282, 0.061)	−0.167 (−0.427, 0.093)	−0.092 (−0.278, 0.093)	−0.140 (−0.421, 0.141)
Sample	2,818		2,766		2,135		1,991	
Friendship satisfaction (not satisfied[a])								
Satisfied	−0.032 (−0.230, 0.165)	−0.049 (−0.347, 0.249)	−0.068 (−0.266, 0.130)	−0.103 (−0402, 0.196)	−0.137 (−0.388, 0.114)	−0.207 (−0.587, 0.173)	−0.079 (−0.348, 0.190)	−0.119 (−0.526, 0.288)
Sample	2,804		2,754		2,127		1,991	
Loneliness (not lonely[a])								
Lonely	−0.010 (−0.063, 0.043)	−0.015 (−0.095, 0.066)	0.004 (−0.051, 0.059)	0.006 (−0.077, 0.089)	0.034 (−0.046, 0.113)	0.051 (−0.069, 0.171)	−0.031 (−0.151, 0.088)	−0.047 (−0.228, 0.133)
Sample	8,145		8,019		4,716		1,991	

Supplemental Table 11 Associations between log errors in the visual memory task and proxies of social engagement. ***Continued.***

Variable	Unadjusted		Partially adjusted		Fully adjusted		Complete	
	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)	Unstd. beta (99.7% CI)	Std. beta (99.7% CI)
Confiding in someone (never[a])								
Monthly	−0.012 (−0.097, 0.073)	−0.018 (−0.147, 0.110)	−0.015 (−0.101, 0.070)	−0.023 (−0.153, 0.106)	−0.005 (−0.116, 0.106)	−0.007 (−0.175, 0.161)	0.003 (−0.168, 0.174)	0.004 (−0.254, 0.263)
Weekly	−0.002 (−0.077, 0.073)	−0.003 (−0.116, 0.110)	−0.002 (−0.078, 0.073)	−0.004 (−0.118, 0.111)	−0.030 (−0.128, 0.068)	−0.046 (−0.194, 0.102)	−0.022 (−0.176, 0.132)	−0.033 (−0.266, 0.199)
Daily	−0.005 (−0.067, 0.056)	−0.008 (−0.101, 0.085)	−0.009 (−0.071, 0.053)	−0.014 (−0.108, 0.080)	−0.027 (−0.106, 0.052)	−0.041 (−0.160, 0.079)	−0.017 (−0.153, 0.118)	−0.026 (−0.231, 0.178)
Sample	7,957		7,837		4,615		1,991	
Social activities (none[a])								
Sports	0.027 (−0.061, 0.116)	0.041 (−0.093, 0.175)	0.027 (−0.062, 0.117)	0.041 (−0.094, 0.176)	−0.023 (−0.132, 0.087)	−0.035 (−0.200, 0.131)	−0.025 (−0.189, 0.138)	−0.038 (−0.286, 0.209)
Pub/social club	−0.010 (−0.078, 0.058)	−0.015 (−0.118, 0.088)	−0.022 (−0.091, 0.047)	−0.033 (−0.138, 0.071)	−0.038 (−0.128, 0.052)	−0.057 (−0.193, 0.079)	0.017 (−0.127, 0.161)	0.026 (−0.193, 0.244)
Religious group	0.034 (−0.059, 0.128)	0.052 (−0.089, 0.194)	0.018 (−0.076, 0.112)	0.027 (−0.116, 0.170)	0.036 (−0.085, 0.157)	0.054 (−0.129, 0.237)	0.058 (−0.124, 0.239)	0.087 (−0.187, 0.362)
Adult education	−0.056 (−0.232, 0.119)	−0.085 (−0.351, 0.180)	−0.050 (−0.225, 0.126)	−0.075 (−0.341, 0.191)	−0.023 (−0.263, 0.217)	−0.035 (−0.397, 0.328)	0.040 (−0.348, 0.427)	0.060 (−0.526, 0.646)
Other	0.034 (−0.049, 0.117)	0.052 (−0.074, 0.177)	0.022 (−0.062, 0.106)	0.033 (−0.093, 0.160)	0.026 (−0.082, 0.134)	0.039 (−0.124, 0.202)	0.079 (−0.088, 0.247)	0.120 (−0.133, 0.373)
Multiple	0.005 (−0.058, 0.068)	0.008 (−0.088, 0.103)	−0.007 (−0.071, 0.057)	−0.010 (−0.107, 0.087)	−0.056 (−0.138, 0.026)	−0.085 (−0.208, 0.039)	−0.024 (−0.149, 0.102)	−0.036 (−0.225, 0.154)
Sample	8,266		8,143		4,778		1,991	

*Significant at $p < 0.003$; [a]reference category

CI indicates confidence interval; std., standardised; unstd., unstandardised.

Supplemental Table 12 Associations between a correct response on the prospective memory task and proxies of social engagement.

	Prospective memory models							
	Unadjusted		Partially adjusted		Fully adjusted		Complete	
Variable	Unstd. beta (99.7% CI)	Odds ratio (99.7% CI)	Unstd. beta (99.7% CI)	Odds ratio (99.7% CI)	Unstd. beta (99.7% CI)	Odds ratio (99.7% CI)	Unstd. beta (99.7% CI)	Odds ratio (99.7% CI)
Family/friend visits (never[a])								
Monthly	−0.072 (−0.757, 0.612)	0.930 (0.469, 1.844)	−0.152 (−0.867, 0.564)	0.859 (0.420, 1.757)	0.116 (−0.747, 0.978)	1.123 (0.474, 2.659)	0.003 (−0.956, 0.961)	1.003 (0.385, 2.613)
Weekly	0.014 (−0.630, 0.658)	1.014 (0.533, 1.931)	−0.079 (−0.754, 0.597)	0.924 (0.471, 1.816)	0.119 (−0.698, 0.937)	1.127 (0.497, 2.552)	−0.033 (−0.948, 0.883)	0.968 (0.387, 2.417)
Daily	−0.135 (−0.829, 0.560)	0.874 (0.436, 1.751)	−0.169 (−0.896, 0.558)	0.844 (0.408, 1.747)	0.092 (−0.789, 0.972)	1.096 (0.454, 2.644)	−0.125 (−1.106, 0.857)	0.883 (0.331, 2.356)
Sample	2,810		2,759		2,131		1,973	
Family relationship satisfaction (not satisfied[a])								
Satisfied	0.088 (−0.330, 0.506)	1.092 (0.719, 1.658)	0.133 (−0.298, 0.564)	1.142 (0.742, 1.758)	−0.049 (−0.616, 0.518)	0.952 (0.540, 1.679)	−0.291 (−0.916, 0.335)	0.748 (0.400, 1.398)
Sample	2,780		2,729		2,114		1,973	
Friendship satisfaction (not satisfied[a])								
Satisfied	0.255 (−0.319, 0.829)	1.290 (0.727, 2.291)	0.323 (−0.265, 0.911)	1.382 (0.767, 2.488)	0.430 (−0.374, 1.234)	1.537 (0.688, 3.435)	0.532 (−0.334, 1.398)	1.702 (0.716, 4.045)
Sample	2,767		2,718		2,106		1,973	
Loneliness (not lonely[a])								
Lonely	−0.364* (−0.631, −0.098)	0.695* (0.532, 0.907)	−0.328* (−0.610, −0.046)	0.720* (0.543, 0.955)	−0.414* (−0.774, −0.054)	0.661* (0.461, 0.947)	−0.305 (−0.695, 0.084)	0.737 (0.499, 1.088)
Sample	2,778		2,727		2,108		1,973	

Supplemental Table 12 Associations between a correct response on the prospective memory task and proxies of social engagement. ***Continued.***

	Unadjusted		Partially adjusted		Fully adjusted		Complete	
Variable	**Unstd. beta (99.7% CI)**	**Odds ratio (99.7% CI)**	**Unstd. beta (99.7% CI)**	**Odds ratio (99.7% CI)**	**Unstd. beta (99.7% CI)**	**Odds ratio (99.7% CI)**	**Unstd. beta (99.7% CI)**	**Odds ratio (99.7% CI)**
Confiding in someone (never[a])								
Monthly	−0.121 (−0.549, 0.307)	0.886 (0.578, 1.359)	−0.219 (−0.660, 0.223)	0.804 (0.517, 1.250)	−0.135 (−0.654, 0.383)	0.873 (0.520, 1.467)	−0.078 (−0.624, 0.467)	0.925 (0.536, 1.596)
Weekly	0.112 (−0.273, 0.498)	1.119 (0.761, 1.646)	0.073 (−0.328, 0.474)	1.076 (0.721, 1.606)	0.133 (−0.340, 0.606)	1.143 (0.712, 1.834)	0.182 (−0.319, 0.684)	1.200 (0.727, 1.981)
Daily	0.233 (−0.102, 0.568)	1.262 (0.903, 1.764)	0.170 (−0.176, 0.516)	1.185 (0.839, 1.675)	0.313 (−0.094, 0.720)	1.367 (0.910, 2.054)	0.343 (−0.096, 0.781)	1.409 (0.909, 2.184)
Sample	2,706		2,660		2,073		1,973	
Social activities (none[a])								
Sports	0.124 (−0.321, 0.569)	1.132 (0.726, 1.766)	0.019 (−0.439, 0.477)	1.019 (0.645, 1.612)	−0.009 (−0.537, 0.519)	0.991 (0.585, 1.681)	−0.016 (−0.574, 0.543)	0.984 (0.563, 1.721)
Pub/social club	−0.172 (−0.522, 0.178)	0.842 (0.593, 1.195)	−0.128 (−0.491, 0.234)	0.880 (0.612, 1.264)	−0.256 (−0.701, 0.188)	0.774 (0.496, 1.207)	−0.275 (−0.744, 0.194)	0.760 (0.475, 1.214)
Religious group	−0.204 (−0.661, 0.254)	0.816 (0.516, 1.289)	−0.191 (−0.663, 0.280)	0.826 (0.515, 1.323)	−0.099 (−0.657, 0.459)	0.906 (0.518, 1.582)	−0.013 (−0.613, 0.587)	0.987 (0.542, 1.798)
Adult education	−0.372 (−1.326, 0.583)	0.690 (0.265, 1.791)	−0.449 (−1.420, 0.522)	0.638 (0.242, 1.685)	−0.175 (−1.319, 0.968)	0.839 (0.268, 2.634)	−0.263 (−1.524, 0.997)	0.769 (0.218, 2.711)
Other	0.118 (−0.327, 0.563)	1.125 (0.721, 1.756)	0.103 (−0.354, 0.561)	1.109 (0.702, 1.752)	0.056 (−0.479, 0.592)	1.058 (0.620, 1.807)	0.104 (−0.461, 0.669)	1.110 (0.630, 1.953)
Multiple	0.241 (−0.086, 0.569)	1.273 (0.917, 1.766)	0.163 (−0.175, 0.501)	1.177 (0.839, 1.650)	0.015 (−0.388, 0.417)	1.015 (0.679, 1.518)	−0.032 (−0.458, 0.394)	0.968 (0.632, 1.482)
Sample	2,815		2,763		2,134		1,973	

*Significant at $p < 0.003$; [a]reference category

CI indicates confidence interval; std., standardised; unstd., unstandardised.

Appendix 6: Chapter 5, correlations between variables included in the model for prediction of acute post-stroke cognitive performance

	Age	Sex	Previous stroke	Previous TIA	Atrial fibrillation	Diabetes	Hypertension	Vascular disease	Stroke severity	Prior dementia	Cognitive performance
Age		0.276*	0.044	0.159*	0.453*	0.047	0.261*	0.233*	0.175*	0.575*	-0.425*
Sex (female)	0.276*		-0.036	0.093	-0.032	-0.018	0.137*	0.060	0.028	0.104	-0.179*
Previous stroke	0.044	-0.036		-0.404*	0.087	0.294*	0.192*	0.197*	0.039	0.260*	-0.088
Previous TIA	0.159*	0.093	-0.404*		-0.068	-0.019	0.018	0.147	-0.237*	-0.068	0.104
Atrial fibrillation	0.453*	-0.032	0.087	-0.068		0.132	0.228*	0.110	0.243*	0.246*	-0.309*
Diabetes	0.047	-0.018	0.294*	-0.019	0.132		0.290*	0.374*	0.079	0.066	-0.087
Hypertension	0.261*	0.137*	0.192*	0.018	0.228*	0.290*		0.295*	0.060	0.035	-0.087
Vascular disease	0.233*	0.060	0.197*	0.147	0.110	0.374*	0.295*		-0.031	0.291*	-0.021
Stroke severity	0.175*	0.028	0.039	-0.237*	0.243*	0.079	0.060	-0.031		0.246*	-0.550*
Prior dementia	0.575*	0.104	0.260*	-0.068	0.246*	0.066	0.035	0.291*	0.246*		-0.609*
Cognitive performance	-0.425*	-0.179*	-0.088	0.104	-0.309*	-0.087	-0.087	-0.021	-0.550*	-0.609	

*Significant at $p < 0.05$

TIA indicates transient ischaemic attack.

Note: Coefficients were obtained from correlations appropriate to each pair of variables: tetrachoric for two dichotomous, biserial for one dichotomous and one continuous, polychoric for two ordered categorical and one ordered categorical and one dichotomous, polyserial for one ordered categorical and one continuous.

Appendix 7: Chapter 5, results of sensitivity analysis

Supplemental Table 13 Direct associations between predictors and stroke severity, dementia and cognitive performance.

	Unstandardised coefficients (95% bias-corrected CI)		
	Stroke severity	**Prior dementia**	**Cognitive performance**
Age	0.012 (0.004, 0.020)*	0.061 (0.037, 0.081)*	0.002 (-0.017, 0.025)
Sex (female)	-0.066 (-0.254, 0.141)	-0.080 (-0.444, 0.296)	-0.188 (-0.526, 0.157)
Previous stroke	0.033 (-0.184, 0.259)	0.396 (-0.053, 0.773)	0.103 (-0.278, 0.517)
Previous TIA	-0.394 (-0.776, 0.006)*	-0.256 (-1.163, 0.305)	-0.012 (-0.727, 0.571)
Atrial fibrillation	0.373 (0.072, 0.641)*	0.053 (-0.427, 0.497)	-0.124 (-0.508, 0.351)
Diabetes	-0.035 (-0.307, 0.246)	0.009 (-0.771, 0.556)	-0.022 (-0.671, 0.522)
Hypertension	0.116 (-0.161, 0.328)	-0.234 (-0.762, 0.270)	-0.081 (-0.554, 0.318)
Vascular disease	0.047 (-0.382, 0.450)	0.644 (-0.016, 1.265)*	0.339 (-0.246, 1.066)
Vascular disease x diabetes	0.456 (-0.024, 0.956)	———	———
Vascular disease x hypertension	-0.482 (-0.983, 0.006)	———	———

*significant at $p < 0.05$

CI indicates confidence interval; TIA, transient ischaemic attack.

Supplemental Table 14 Indirect associations between predictors and cognitive performance.

	Unstandardised coefficients (95% bias-corrected CI)	
	Effects mediated through stroke severity	**Effects mediated through prior dementia**
Age	-0.008 (-0.015, -0.003)*	-0.041 (-0.069, -0.021)*
Sex (female)	0.048 (-0.104, 0.197)	0.054 (-0.214, 0.307)
Previous stroke	-0.024 (-0.195, 0.138)	-0.265 (-0.599, 0.001)
Previous TIA	0.285 (-0.020, 0.584)	0.171 (-0.217, 0.826)
Atrial fibrillation	-0.269 (-0.492, -0.049)*	-0.035 (-0.387, 0.270)
Diabetes	0.025 (-0.195, 0.234)	-0.006 (-0.393, 0.542)
Hypertension	-0.084 (-0.263, 0.110)	0.157 (-0.165, 0.588)
Vascular disease	-0.034 (-0.348, 0.290)	-0.432 (-1.017, -0.023)
Vascular disease x diabetes	-0.330 (-0.741, 0.022)	———
Vascular disease x hypertension	0.348 (-0.011, 0.748)	———

*significant at $p < 0.05$

CI indicates confidence interval; TIA, transient ischaemic attack.

Appendix 8: Chapter 6, case report form for recording baseline clinical information in APPLE

APPLE Study

Protocol Version 1.5

Version 4.0 (26 Feb 2018)

Baseline Assessment

Informed Consent and Participant Characteristics

Site Number	Participant Number	Initials	Date of Visit
01	02	03	D D M M Y Y Y Y 04

A. Informed Consent

1. Does the participant have the capacity to provide written, witnessed informed consent for study participation? Yes ☐ 1 No ☐ 2 (05)

If Yes, has written, witnessed informed consent for study participation been provided? Yes ☐ 1 No ☐ 2 (06)

If Yes, has written, witnessed informed consent for additional study A been provided? Yes ☐ 1 No ☐ 2 (07)

If Yes, has written, witnessed informed consent for additional study B been provided *(NHS GG&C only - other sites select N/A)*? Yes ☐ 1 No ☐ 2 N/A ☐ 3 (08)

If Yes, date consent signed D D M M Y Y Y Y 09

If No, has an appropriate proxy provided written, witnessed informed consent for study participation? Yes ☐ 1 No ☐ 2 (10)

If No, has an appropriate proxy provided written, witnessed informed consent for additional study A participation? Yes ☐ 1 No ☐ 2 (11)

If No, has an appropriate proxy provided written, witnessed informed consent for additional study B participation *(NHS GG&C only - other sites select N/A)*? Yes ☐ 1 No ☐ 2 N/A ☐ 3 (12)

If Yes, date consent signed D D M M Y Y Y Y 13

2. Has written, witnessed informed consent for study participation been provided by the partcipant's preferred carer/friend/relative (Study Informant)? Yes ☐ 1 No ☐ 2 (14)

If Yes, date consent signed D D M M Y Y Y Y15

To minimise test burden and patient interview time, most of the baseline information should be available from medical and nursing notes, Parts D and E may require direct patient questioning.

B. Participant Characteristics

1. Date of Birth D D M M Y Y Y Y 16
2. Gender Male ☐ 1 Female ☐ 2 (17)
3. Height (cm) ___.___ (18)
4. Weight (kg) ___.___ (19)

C. For NHS Greater Glasgow & Clyde sites only

1. Research Database ID (20)

Produced by Robertson Centre for Biostatistics, University of Glasgow

APPLE Study

Protocol Version 1.3

Version 3.0 (22 Jun 2017)

Baseline Assessment
Inclusion and Exclusion Criteria

Site Number	Participant Number	Initials	Date of Visit
01	02	03	D D M M Y Y Y Y 04

A. Inclusion Criteria

1. Informed consent has been provided by the participant / proxy — Yes ☐ 1 No* ☐ 2 (05)
2. Clinical diagnosis of stroke or transient ischaemic attack (TIA) at time of assessment. — Yes ☐ 1 No* ☐ 2 (06)
3. Age greater than 18 years — Yes ☐ 1 No* ☐ 2 (07)
4. Treating clinician happy that the patient would have some form of psychological assessment as part of usual care. — Yes ☐ 1 No* ☐ 2 (08)

***IF THE ANSWER IS NO TO ANY OF THE ABOVE, THE PARTICIPANT IS NOT ELIGIBLE TO CONTINUE IN THE STUDY**

B. Exclusion Criteria

1. Non-stroke diagnosis at time of assessment — Yes* ☐ 1 No ☐ 2 (09)
2. Unable to consent and no suitable proxy available — Yes* ☐ 1 No ☐ 2 (10)
3. No spoken English pre-stroke — Yes* ☐ 1 No ☐ 2 (11)
4. Prisoners — Yes* ☐ 1 No ☐ 2 (12)

***IF THE ANSWER IS YES TO ANY OF THE ABOVE, THE PARTICIPANT IS NOT ELIGIBLE TO CONTINUE IN THE STUDY**

APPLE Study

Protocol Version 1.3

Version 3.0 (22 Jun 2017)

Baseline Assessment
Medical History

Site Number	Participant Number	Initials	Date of Visit
01	02	03	D D M M Y Y Y Y 04

A. Stroke/TIA details

1. a) Date / time of stroke/ TIA — D D M M Y Y Y Y 05 — H H M M 06
   b) Date / time of stroke unit admission — D D M M Y Y Y Y 07 — H H M M 08
   c) Date / time of assessment — D D M M Y Y Y Y 09 — H H M M 10

B. Stroke Risk Factors

1. Atrial fibrillation — Yes ☐ 1 No ☐ 2 (11)
2. Diabetes mellitus — Yes ☐ 1 No ☐ 2 (12)
3. Heart Failure — Yes ☐ 1 No ☐ 2 (13)
4. Hypertension — Yes ☐ 1 No ☐ 2 (14)
   If yes, Treated ☐ 1 Untreated ☐ 2 (15)
5. Hyperlipidaemia — Yes ☐ 1 No ☐ 2 (16)
6. Migraine — Yes ☐ 1 No ☐ 2 (17)
7. Previous stroke/TIA — Yes ☐ 1 No ☐ 2 (18)
   If yes, M M Y Y Y Y 19
8. Seizure disorder — Yes ☐ 1 No ☐ 2 (20)
9. Vascular Disease — Yes ☐ 1 No ☐ 2 (21)

C. Past Medical History

1. COPD — Yes ☐ 1 No ☐ 2 (22)
2. Connective tissue disease (rheumatoid, lupus) — Yes ☐ 1 No ☐ 2 (23)
3. Peptic ulcer — Yes ☐ 1 No ☐ 2 (24)
4. Liver disease — Yes ☐ 1 No ☐ 2 (25)
5. Renal disease — Yes ☐ 1 No ☐ 2 (26)
6. Tumour no-metastasis — Yes ☐ 1 No ☐ 2 (27)
7. Metastatic tumour — Yes ☐ 1 No ☐ 2 (28)

APPLE Study

Protocol Version 1.3

Version 3.0 (22 Jun 2017)

Baseline Assessment
Medical History

Site Number	Participant Number	Initials	Date of Visit
01	02	03	D D M M Y Y Y Y 04

D. Psychological History

1. Any history previous mood disorder requiring treatment — Yes ☐ 1 No ☐ 2 (05)
2. Any history previous delirium — Yes ☐ 1 No ☐ 2 (06)
3. Any history dementia or cognitive impairment — Yes ☐ 1 No ☐ 2 (07)
4. Any history alcohol or substance dependence — Yes ☐ 1 No ☐ 2 (08)
5. Years in education — ☐☐ (09)
6. English first language — Yes ☐ 1 No ☐ 2 (10)

E. Family History

1. Premature vascular disease in any first degree relative(age <55 years) — Yes ☐ 1 No ☐ 2 Unknown ☐ 3 (11)
2. History of dementia in mother — Yes ☐ 1 No ☐ 2 Unknown ☐ 3 (12)
   - **If Yes,** give age of onset ☐☐ (13)
3. History of dementia in father — Yes ☐ 1 No ☐ 2 Unknown ☐ 3 (14)
   - **If Yes,** give age of onset ☐☐ (15)

APPLE Study

Protocol Version 1.3

Version 3.0 (22 Jun 2017)

Site Number	Participant Number	Initials	Date of Visit
01	02	03	D D M M Y Y Y Y 04

A. Social and Functional Assessment

1. Smoking — Current ☐1 Former ☐2 Never ☐3 (05)
   If current or former: a) number of pack years (06)
2. Alcohol — Yes ☐1 No ☐2 (07)
   If yes: a) units per week (08)
3. Care-home resident — Yes ☐1 No ☐2 (09)
4. Continent bowel — Yes ☐1 No ☐2 (10)
5. Continent bladder — Yes ☐1 No ☐2 (11)
6. Number or carers (daily) (12)
7. Mobility — Independent ☐1 Needs walking aid ☐2 Needs assistance ☐3 (13)
8. History of > 2 falls in last year — Yes ☐1 No ☐2 (14)
9. Visual impairment — Yes ☐1 No ☐2 (15)
10. Hearing impairment — Yes ☐1 No ☐2 (16)

B. Examination (where available take this information from the last observations recorded)

1. Date / time of observations used — D D M M Y Y Y Y (17) H H M M (18)
2. Dominant hand — Left ☐1 Right ☐2 Ambidextrous ☐3 (19)
3. Glasgow Coma Scale — E ☐1 M ☐2 V ☐3 (20)
4. AVPU — alert ☐1 responds verbal stimuli ☐2 responds painful stimuli ☐3 unresponsive ☐4 (21)
5. Temperature (22) °C
6. Blood glucose (23) mmol/L
7. Heart rate (24) bpm
8. Any atrial fibrillation documented in this admission" — Yes ☐1 No ☐2 (25)
9. Swallow — safe ☐1 unsafe ☐2 not tested ☐3 (26)
10. Blood pressure — SBP (27) DBP (28) mmhg

APPLE Study

Protocol Version 1.3

Version 3.0 (22 Jun 2017)

Baseline Assessment
Medication, Stroke Details and Blood Results
Page 6 of 6

Site Number	Participant Number	Initials	Date of Visit
01	02	03	D D M M Y Y Y Y 04

A. Medications

1. Is the participant currently taking any medications? Yes ☐ 1 No ☐ 2 05

If Yes, please enter in Concomitant Therapy Log

B. Stroke Classification

tick one, if bilateral or two territories tick all that apply"

1. TACS — Left ☐ 1 Right ☐ 2 06
2. PACS — Left ☐ 1 Right ☐ 2 07
3. LACS — Left ☐ 1 Right ☐ 2 08
4. POCS — Yes ☐ 1 No ☐ 2 09
5. TIA — anterior left ☐ 1 anterior right ☐ 2 posterior ☐ 3 10
6. Unknown — Yes ☐ 1 No ☐ 2 11

C. Blood Results

1. Sodium ___ mmol/L 12
2. Urea ___.___ mmol/L 13
3. Creatinine ___ umol/L 14
4. Haemoglobin ___ g/dL 15
5. White cell count ___.___ x10 ^{9}/l 16
6. C Reactive Protein ___ mg/L 17
7. Albumin ___ g/L 18
8. Calcium (adjusted) ___.___ mmol/L 19

FOR OFFICE USE ONLY (not for data entry)

1. Arrange a date for the semi-structured clinical interview (within one month of baseline visit) — Date of Interview D D M M Y Y Y Y

2. Arrange a date for the next study visit (1 Month Follow-up) — Date of Interview D D M M Y Y Y Y

Produced by Robertson Centre for Biostatistics, University of Glasgow

Appendix 9: Chapter 7, Mplus code for factor analysis models

Configural model

```
TITLE: CFA for AMT data baseline to 12m, configural model;

DATA:  FILE IS
C:\Users\Bogna\OneDrive - University of Glasgow\APPLE Trajectories\AMT_raw
data.dat;

VARIABLE:

    NAMES ARE

       ID_new
       ori_B ww1_B att_B rec_B clock_B news_B let_B
       ori_1 ww1_1 att_1 rec_1 clock_1 news_1 let_1
       ori_6 ww1_6 att_6 rec_6 clock_6 news_6 let_6
       ori_12 ww1_12 att_12 rec_12 clock_12 news_12 let_12;

    USEVARIABLES  ARE

       ori_B ww1_B att_B rec_B clock_B news_B let_B
       ori_1 ww1_1 att_1 rec_1 clock_1 news_1 let_1
       ori_6 ww1_6 att_6 rec_6 clock_6 news_6 let_6
       ori_12 ww1_12 att_12 rec_12 clock_12 news_12 let_12;

  CATEGORICAL ARE ALL;

  IDVARIABLE IS ID_new;

  MISSING ARE ALL (-999);

ANALYSIS:

  ESTIMATOR = WLSMV;

  PROCESSORS = 2;

MODEL:

  AMT0 BY ori_B ww1_B att_B rec_B clock_B news_B let_B;
  AMT1 BY ori_1 ww1_1 att_1 rec_1 clock_1 news_1 let_1;
  AMT2 BY ori_6 ww1_6 att_6 rec_6 clock_6 news_6 let_6;
  AMT3 BY ori_12 ww1_12 att_12 rec_12 clock_12 news_12 let_12;

  {ori_B@1 ww1_B@1 att_B@1 rec_B@1 clock_B@1 news_B@1 let_B@1
  ori_1@1 ww1_1@1 att_1@1 rec_1@1 clock_1@1 news_1@1 let_1@1
  ori_6@1 ww1_6@1 att_6@1 rec_6@1 clock_6@1 news_6@1 let_6@1
  ori_12@1 ww1_12@1 att_12@1 rec_12@1 clock_12@1 news_12@1 let_12@1};

  [AMT0@0 AMT1@0 AMT2@0 AMT3@0];

  AMT0 WITH AMT1 AMT2 AMT3;
  AMT1 WITH AMT2 AMT3;
  AMT2 WITH AMT3;
```

```
   ori_B WITH ori_1 ori_6 ori_12;
   ori_1 WITH ori_6 ori_12;
   ori_6 WITH ori_12;

   ww1_B WITH ww1_1 ww1_6 ww1_12;
   ww1_1 WITH ww1_6 ww1_12;
   ww1_6 WITH ww1_12;

   att_B WITH att_1 att_6 att_12;
   att_1 WITH att_6 att_12;
   att_6 WITH att_12;

   rec_B WITH rec_1 rec_6 rec_12;
   rec_1 WITH rec_6 rec_12;
   rec_6 WITH rec_12;

   clock_B WITH clock_1 clock_6 clock_12;
   clock_1 WITH clock_6 clock_12;
   clock_6 WITH clock_12;

   news_B WITH news_1 news_6 news_12;
   news_1 WITH news_6 news_12;
   news_6 WITH news_12;

   let_B WITH let_1 let_6 let_12;
   let_1 WITH let_6 let_12;
   let_6 WITH let_12;
OUTPUT:
   STANDARDIZED MODINDICES(3.84) PATTERNS RESIDUAL;
SAVEDATA:
   DIFFTEST = chidiff.dat;
```

Scalar model

```
TITLE: CFA for AMT data baseline to 12m, scalar model;
DATA:  FILE IS
C:\Users\Bogna\OneDrive - University of Glasgow\APPLE Trajectories\AMT_raw
data.dat;
VARIABLE:
   NAMES ARE
      ID_new
      ori_B ww1_B att_B rec_B clock_B news_B let_B
      ori_1 ww1_1 att_1 rec_1 clock_1 news_1 let_1
      ori_6 ww1_6 att_6 rec_6 clock_6 news_6 let_6
      ori_12 ww1_12 att_12 rec_12 clock_12 news_12 let_12;
  USEVARIABLES  ARE
      ori_B ww1_B att_B rec_B clock_B news_B let_B
      ori_1 ww1_1 att_1 rec_1 clock_1 news_1 let_1
      ori_6 ww1_6 att_6 rec_6 clock_6 news_6 let_6
      ori_12 ww1_12 att_12 rec_12 clock_12 news_12 let_12;
```

```
    CATEGORICAL ARE ALL;
    IDVARIABLE IS ID_new;
    MISSING ARE ALL (-999);
ANALYSIS:
    ESTIMATOR = WLSMV;
    PROCESSORS = 2;
DIFFTEST = C:\Users\Bogna\OneDrive - University of Glasgow\APPLE
Trajectories\chidiff.dat;
MODEL:
    AMT0 BY ori_B;
    AMT0 BY ww1_B (cow);
    AMT0 BY att_B (hen);
    AMT0 BY rec_B (horse);
    AMT0 BY clock_B (donkey);
    AMT0 BY news_B (deer);
    AMT0 BY let_B (elk);
    AMT1 BY ori_1;
    AMT1 BY ww1_1 (cow);
    AMT1 BY att_1 (hen);
    AMT1 BY rec_1 (horse);
    AMT1 BY clock_1 (donkey);
    AMT1 BY news_1 (deer);
    AMT1 BY let_1 (elk);
    AMT2 BY ori_6;
    AMT2 BY ww1_6 (cow);
    AMT2 BY att_6 (hen);
    AMT2 BY rec_6 (horse);
    AMT2 BY clock_6 (donkey);
    AMT2 BY news_6 (deer);
    AMT2 BY let_6 (elk);
    AMT3 BY ori_12;
    AMT3 BY ww1_12 (cow);
    AMT3 BY att_12 (hen);
    AMT3 BY rec_12 (horse);
    AMT3 BY clock_12 (donkey);
    AMT3 BY news_12 (deer);
    AMT3 BY let_12 (elk);
    [ori_B$1] (yellow);
    [ww1_B$1] (black);
    [att_B$1] (indigo);
    [rec_B$1] (ochre);
    [clock_B$2] (turquoise);
    [news_B$1] (gold);
    [let_B$1] (magenta);
    [ori_1$1] (yellow);
    [ww1_1$1] (black);
    [att_1$1] (indigo);
```

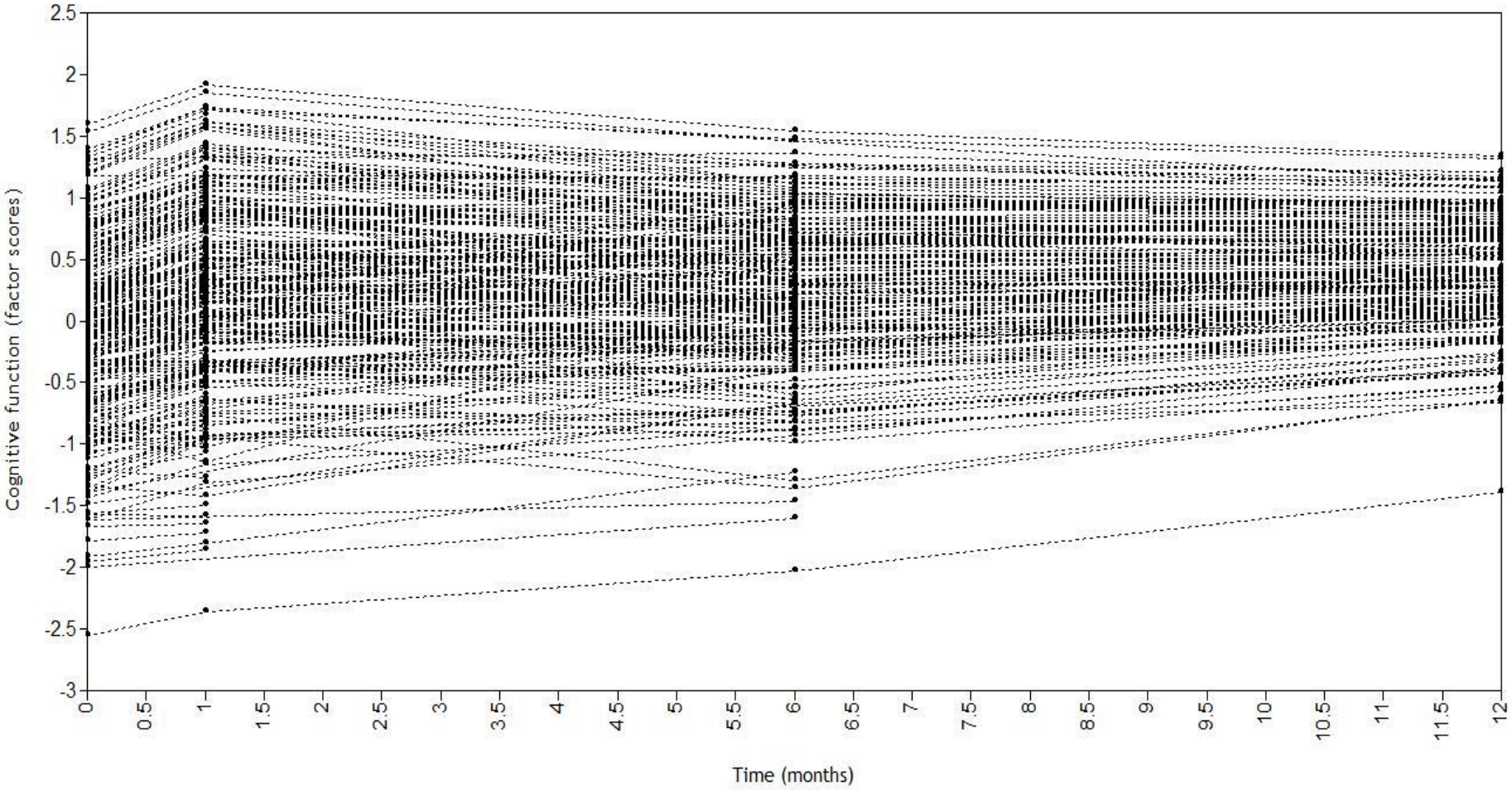
```
[rec_1$1] (ochre);
[clock_1$2] (turquoise);
[news_1$1] (gold);
[let_1$1] (magenta);

[ori_6$1] (yellow);
[ww1_6$1] (black);
[att_6$1] (indigo);
[rec_6$1] (ochre);
[clock_6$2] (turquoise);
[news_6$1] (gold);
[let_6$1] (magenta);

[ori_12$1] (yellow);
[ww1_12$1] (black);
[att_12$1] (indigo);
[rec_12$1] (ochre);
[clock_12$2] (turquoise);
[news_12$1] (gold);
[let_12$1] (magenta);

{ori_B@1 ww1_B@1 att_B@1 rec_B@1 clock_B@1 news_B@1 let_B@1
ori_1 ww1_1 att_1 rec_1 clock_1 news_1 let_1
ori_6 ww1_6 att_6 rec_6 clock_6 news_6 let_6

ori_12 ww1_12 att_12 rec_12 clock_12 news_12 let_12};

[AMT0@0 AMT1* AMT2* AMT3*];

AMT0 WITH AMT1 AMT2 AMT3;
AMT1 WITH AMT2 AMT3;
AMT2 WITH AMT3;

ori_B WITH ori_1 ori_6 ori_12;
ori_1 WITH ori_6 ori_12;
ori_6 WITH ori_12;

ww1_B WITH ww1_1 ww1_6 ww1_12;
ww1_1 WITH ww1_6 ww1_12;
ww1_6 WITH ww1_12;

att_B WITH att_1 att_6 att_12;
att_1 WITH att_6 att_12;
att_6 WITH att_12;

rec_B WITH rec_1 rec_6 rec_12;
rec_1 WITH rec_6 rec_12;
rec_6 WITH rec_12;

clock_B WITH clock_1 clock_6 clock_12;
clock_1 WITH clock_6 clock_12;
clock_6 WITH clock_12;

news_B WITH news_1 news_6 news_12;
news_1 WITH news_6 news_12;
news_6 WITH news_12;

let_B WITH let_1 let_6 let_12;
let_1 WITH let_6 let_12;
let_6 WITH let_12;
```

```
OUTPUT:
   STANDARDIZED MODINDICES(3.84) PATTERNS;
SAVEDATA:
   SAVE=FS;
   FILE IS AMT_fscores_scalar.dat;
```

Appendix 10: Chapter 7, Mplus code for a latent class trajectory model with three classes

```
TITLE: LCGA with default approach to handling missing outcome data;
DATA: FILE IS
C:\Users\Bogna\OneDrive - University of Glasgow\APPLE Trajectories\
\LCGA_default\AMT_FS&T_finalN_def.dat;
VARIABLE:
   NAMES ARE
      ID_new
      AMT0 AMT1 AMT2 AMT3 t0 t1 t2 t3;
   USEVARIABLES ARE
      AMT0 AMT1 AMT2 AMT3 t0 t1 t2 t3;
   TSCORES = t0 t1 t2 t3;
   MISSING ARE ALL (-999);
   IDVARIABLE IS ID_new;
   CLASSES = c(3);
ANALYSIS:
   TYPE = MIXTURE RANDOM;
   ALGORITHM = INTEGRATION;
   STARTS = 100 10;
   STITERATIONS = 10;
   PROCESSORS = 4;
MODEL:
   %OVERALL%
   i s | AMT0-AMT3 AT t0-t3;
   i@0 s@0;
PLOT:
   TYPE = PLOT3;
   SERIES = AMT0(0) AMT1(1) AMT2(6) AMT3(12);
OUTPUT:
   sampstat TECH1 TECH8;
```

Appendix 11: Chapter 8, individual trajectories of change in cognitive function over a 12-month period, following recruitment into the APPLE study.



List of References

1. Horn JL, Cattell RB. Refinement and test of the theory of fluid and crystallized general intelligences. *J Educ Psychol* (1966) 57(5):253-70. Epub 1966/10/01. doi: 10.1037/h0023816. PubMed PMID: 5918295.

2. Salthouse T. Consequences of age-related cognitive declines. *Annu Rev Psychol* (2012) 63:201-26. Epub 2011/07/12. doi: 10.1146/annurev-psych-120710-100328. PubMed PMID: 21740223.

3. Harada CN, Natelson Love MC, Triebel KL. Normal cognitive aging. *Clin Geriatr Med* (2013) 29(4):737-52. Epub 2013/10/08. doi: 10.1016/j.cger.2013.07.002. PubMed PMID: 24094294.

4. Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, Ivnik RJ, et al. Mild cognitive impairment: ten years later. *Arch Neurol* (2009) 66(12):1447-55. Epub 2009/12/17. doi: 10.1001/archneurol.2009.266. PubMed PMID: 20008648; PubMed Central PMCID: PMCPMC3081688.

5. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers & Dementia* (2011) 7(3):270-9. doi: 10.1016/j.jalz.2011.03.008. PubMed PMID: WOS:000291239600004.

6. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (5th ed.)*. Washington, DC: American Psychiatric Association (2013).

7. Sachs-Ericsson N, Blazer DG. The new DSM-5 diagnosis of mild neurocognitive disorder and its relation to research in mild cognitive impairment. *Aging Ment Health* (2015) 19(1):2-12. Epub 2014/06/11. doi: 10.1080/13607863.2014.920303. PubMed PMID: 24914889.

8. Sachdev PS, Blacker D, Blazer DG, Ganguli M, Jeste DV, Paulsen JS, et al. Classifying neurocognitive disorders: the DSM-5 approach. *Nat Rev Neurol* (2014) 10(11):634-42. Epub 2014/10/01. doi: 10.1038/nrneurol.2014.181. PubMed PMID: 25266297.

9. Deckers K, van Boxtel MP, Schiepers OJ, de Vugt M, Munoz Sanchez JL, Anstey KJ, et al. Target risk factors for dementia prevention: a systematic review and Delphi consensus study on the evidence from observational studies. *Int J Geriatr Psychiatry* (2015) 30(3):234-46. Epub 2014/12/17. doi: 10.1002/gps.4245. PubMed PMID: 25504093.

10. Anstey KJ, Ee N, Eramudugolla R, Jagger C, Peters R. A Systematic Review of Meta-Analyses that Evaluate Risk Factors for Dementia to Evaluate the Quantity, Quality, and Global Representativeness of Evidence. *J Alzheimers Dis* (2019) 70:S165-S86. doi: 10.3233/Jad-190181. PubMed PMID: WOS:000481670800014.

11. Peters R, Booth A, Rockwood K, Peters J, D'Este C, Anstey KJ. Combining modifiable risk factors and risk of dementia: a systematic review and meta-analysis. *Bmj Open* (2019) 9(1). doi: ARTN e022846 10.1136/bmjopen-2018-022846. PubMed PMID: WOS:000471116800067.

12. Johnson CO, Nguyen M, Roth GA, Nichols E, Alam T, Abate D, et al. Global, regional, and national burden of stroke, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* (2019) 18(5):439-58. doi: 10.1016/S1474-4422(19)30034-1. PubMed PMID: WOS:000464140400013.

13. Leys D, Henon H, Mackowiak-Cordoliani MA, Pasquier F. Poststroke dementia. *Lancet Neurol* (2005) 4(11):752-9. Epub 2005/10/22. doi: 10.1016/S1474-4422(05)70221-0. PubMed PMID: 16239182.

14. WHO MONICA Project Principal Investigators. The World Health Organization MONICA Project (monitoring trends and determinants in cardiovascular disease): a major international collaboration. . *J Clin Epidemiol* (1988) 41(2):105-14. Epub 1988/01/01. doi: 10.1016/0895-4356(88)90084-4. PubMed PMID: 3335877.

15. National Institute for Health and Care Excellence (NICE). What are strokes and transient ischaemic attacks? : NICE [updated August 2020; cited 2021 March 2]. Available from: https://cks.nice.org.uk/topics/stroke-tia/background-information/definition/.

16. Sacco RL, Kasner SE, Broderick JP, Caplan LR, Connors JJ, Culebras A, et al. An updated definition of stroke for the 21st century: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* (2013) 44(7):2064-89. Epub 2013/05/09. doi: 10.1161/STR.0b013e318296aeca. PubMed PMID: 23652265.

17. Murphy SJ, Werring DJ. Stroke: causes and clinical features. *Medicine (Abingdon)* (2020) 48(9):561-6. Epub 2020/08/25. doi: 10.1016/j.mpmed.2020.06.002. PubMed PMID: 32837228.

18. Easton JD, Saver JL, Albers GW, Alberts MJ, Chaturvedi S, Feldmann E, et al. Definition and evaluation of transient ischemic attack: a scientific statement for healthcare professionals from the American Heart Association/American Stroke Association Stroke Council; Council on Cardiovascular Surgery and Anesthesia; Council on Cardiovascular Radiology and Intervention; Council on Cardiovascular Nursing; and the Interdisciplinary Council on Peripheral Vascular Disease. The American Academy of Neurology affirms the value of this statement as an educational tool for neurologists. *Stroke* (2009) 40(6):2276-93. Epub 2009/05/09. doi: 10.1161/STROKEAHA.108.192218. PubMed PMID: 19423857.

19. van Rooij FG, Kessels RP, Richard E, De Leeuw FE, van Dijk EJ. Cognitive Impairment in Transient Ischemic Attack Patients: A Systematic Review. *Cerebrovasc Dis* (2016) 42(1-2):1-9. Epub 2016/02/18. doi: 10.1159/000444282. PubMed PMID: 26886189.

20. Pendlebury ST, Rothwell PM, Oxford Vascular S. Incidence and prevalence of dementia associated with transient ischaemic attack and stroke: analysis of the population-based Oxford Vascular Study. *Lancet Neurol* (2019) 18(3):248-58. Epub 2019/02/21. doi: 10.1016/S1474-4422(18)30442-3. PubMed PMID: 30784556.

21. Shaw RC, Walker G, Elliott E, Quinn TJ. Occurrence Rate of Delirium in Acute Stroke Settings Systematic Review and Meta-Analysis. *Stroke* (2019) 50(11):3028-36. doi: 10.1161/Strokeaha.119.025015. PubMed PMID: WOS:000492999300032.

22. Fong TG, Tulebaev SR, Inouye SK. Delirium in elderly adults: diagnosis, prevention and treatment. *Nature Reviews Neurology* (2009) 5(4):210-20. doi: 10.1038/nrneurol.2009.24. PubMed PMID: WOS:000265854600008.

23. Pendlebury ST, Rothwell PM. Prevalence, incidence, and factors associated with pre-stroke and post-stroke dementia: a systematic review and meta-analysis. *Lancet Neurol* (2009) 8(11):1006-18. Epub 2009/09/29. doi: 10.1016/S1474-4422(09)70236-4. PubMed PMID: 19782001.

24. Melkas S, Jokinen H, Hietanen M, Erkinjuntti T. Poststroke cognitive impairment and dementia: prevalence, diagnosis, and treatment. *Degener Neurol Neuromuscul Dis* (2014) 4:21-7. Epub 2014/02/05. doi: 10.2147/DNND.S37353. PubMed PMID: 32669898.

25. Sun JH, Tan L, Yu JT. Post-stroke cognitive impairment: epidemiology, mechanisms and management. *Ann Transl Med* (2014) 2(8):80. Epub 2014/10/22. doi: 10.3978/j.issn.2305-5839.2014.08.05. PubMed PMID: 25333055.

26. van Rijsbergen MW, Mark RE, de Kort PL, Sitskoorn MM. Subjective cognitive complaints after stroke: a systematic review. *J Stroke Cerebrovasc Dis* (2014) 23(3):408-20. Epub 2013/06/27. doi: 10.1016/j.jstrokecerebrovasdis.2013.05.003. PubMed PMID: 23800498.

27. Lo JW, Crawford JD, Desmond DW, Godefroy O, Jokinen H, Mahinrad S, et al. Profile of and risk factors for poststroke cognitive impairment in diverse ethnoregional groups. *Neurology* (2019) 93(24):e2257-e71. Epub 2019/11/13. doi: 10.1212/WNL.0000000000008612. PubMed PMID: 31712368.

28. Heruti RJ, Lusky A, Dankner R, Ring H, Dolgopiat M, Barell V, et al. Rehabilitation outcome of elderly patients after a first stroke: Effect of cognitive status at admission on the functional outcome. *Arch Phys Med Rehab* (2002) 83(6):742-9. doi: 10.1053/apmr.2002.32739. PubMed PMID: WOS:000176004400002.

29. Patel MD, Coshall C, Rudd AG, Wolfe CD. Cognitive impairment after stroke: clinical determinants and its associations with long-term stroke outcomes. *J Am Geriatr Soc* (2002) 50(4):700-6. Epub 2002/05/02. doi: 10.1046/j.1532-5415.2002.50165.x. PubMed PMID: 11982671.

30. Nys GMS, van Zandvoort MJE, van der Worp HB, de Haan EHF, de Kort PLM, Jansen BPW, et al. Early cognitive impairment predicts long-term depressive symptoms and quality of life after stroke. *Journal of the Neurological Sciences* (2006) 247(2):149-56. doi: 10.1016/j.jns.2006.04.005.

31. Cumming TB, Brodtmann A, Darby D, Bernhardt J. The importance of cognition to quality of life after stroke. *Journal of Psychosomatic Research* (2014) 77(5):374-9. doi: 10.1016/j.jpsychores.2014.08.009. PubMed PMID: WOS:000344718200006.

32. Mole JA, Demeyere N. The relationship between early post-stroke cognition and longer term activities and participation: A systematic review. *Neuropsychological Rehabilitation* (2020) 30(2):346-70. doi: 10.1080/09602011.2018.1464934. PubMed PMID: WOS:000516877900012.

33. Rigby H, Gubitz G, Phillips S. A systematic review of caregiver burden following stroke. *Int J Stroke* (2009) 4(4):285-92. Epub 2009/08/20. doi: 10.1111/j.1747-4949.2009.00289.x. PubMed PMID: 19689757.

34. Rohde D, Gaynor E, Large M, Conway O, Bennett K, Williams DJ, et al. Stroke survivor cognitive decline and psychological wellbeing of family caregivers five years post-stroke: a cross-sectional analysis. *Top Stroke Rehabil* (2019) 26(3):180-6. Epub 2019/03/26. doi: 10.1080/10749357.2019.1590972. PubMed PMID: 30907273.

35. Tang EYH, Price C, Stephan BCM, Robinson L, Exley C. Impact of Memory Problems Post-stroke on Patients and Their Family Carers: A Qualitative Study. *Front Med-Lausanne* (2020) 7. doi: ARTN 267 10.3389/fmed.2020.00267. PubMed PMID: WOS:000548401400001.

36. Claesson L, Linden T, Skoog I, Blomstrand C. Cognitive impairment after stroke - impact on activities of daily living and costs of care for elderly people. The Goteborg 70+ Stroke Study. *Cerebrovasc Dis* (2005) 19(2):102-9. Epub 2004/12/21. doi: 10.1159/000082787. PubMed PMID: 15608434.

37. Pollock A, St George B, Fenton M, Firkins L. Top ten research priorities relating to life after stroke. *Lancet Neurol* (2012) 11(3):209-. doi: Doi 10.1016/S1474-4422(12)70029-7. PubMed PMID: WOS:000301014900008.

38. Mijajlovic MD, Pavlovic A, Brainin M, Heiss WD, Quinn TJ, Ihle-Hansen HB, et al. Post-stroke dementia - a comprehensive review. *BMC Med* (2017) 15(1):11. Epub 2017/01/18. doi: 10.1186/s12916-017-0779-7. PubMed PMID: 28095900,

39. Brainin M, Tuomilehto J, Heiss WD, Bornstein NM, Bath PM, Teuschl Y, et al. Post-stroke cognitive decline: an update and perspectives for clinical research. *Eur J Neurol* (2015) 22(2):229-38, e13-6. Epub 2014/12/11. doi: 10.1111/ene.12626. PubMed PMID: 25492161.

40. Broome LJ, Battle CE, Lawrence M, Evans PA, Dennis MS. Cognitive Outcomes following Thrombolysis in Acute Ischemic Stroke: A Systematic Review. *J Stroke Cerebrovasc Dis* (2016) 25(12):2868-75. Epub 2016/08/30. doi: 10.1016/j.jstrokecerebrovasdis.2016.07.048. PubMed PMID: 27569709.

41. Strambo D, Bartolini B, Beaud V, Marto JP, Sirimarco G, Dunet V, et al. Thrombectomy and Thrombolysis of Isolated Posterior Cerebral Artery Occlusion: Cognitive, Visual, and Disability Outcomes. *Stroke* (2020) 51(1):254-61. Epub 2019/11/14. doi: 10.1161/STROKEAHA.119.026907. PubMed PMID: 31718503.

42. Bath PM, Scutt P, Blackburn DJ, Ankolekar S, Krishnan K, Ballard C, et al. Intensive versus Guideline Blood Pressure and Lipid Lowering in Patients with Previous Stroke: Main Results from the Pilot 'Prevention of Decline in Cognition after Stroke Trial' (PODCAST) Randomised Controlled Trial. *PLoS One* (2017) 12(1):e0164608. Epub 2017/01/18. doi: 10.1371/journal.pone.0164608. PubMed PMID: 28095412.

43. Tzourio C, Anderson C, Chapman N, Woodward M, Neal B, MacMahon S, et al. Effects of blood pressure lowering with perindopril and indapamide therapy on dementia and cognitive decline in patients with cerebrovascular disease. *Arch Intern Med* (2003) 163(9):1069-75. Epub 2003/05/14. doi: 10.1001/archinte.163.9.1069. PubMed PMID: 12742805.

44. Pearce LA, McClure LA, Anderson DC, Jacova C, Sharma M, Hart RG, et al. Effects of long-term blood pressure lowering and dual antiplatelet treatment on cognitive function in patients with recent lacunar stroke: a secondary analysis from the SPS3 randomised trial. *Lancet Neurol* (2014) 13(12):1177-85. Epub 2014/12/03. doi: 10.1016/S1474-4422(14)70224-8. PubMed PMID: 25453457.

45. Ankolekar S, Geeganage C, Anderton P, Hogg C, Bath PM. Clinical trials for preventing post stroke cognitive impairment. *J Neurol Sci* (2010) 299(1-2):168-74. Epub 2010/09/22. doi: 10.1016/j.jns.2010.08.052. PubMed PMID: 20855090.

46. Yang Z, Wang H, Edwards D, Ding C, Yan L, Brayne C, et al. Association of blood lipids, atherosclerosis and statin use with dementia and cognitive impairment after stroke: A systematic review and meta-analysis. *Ageing Res Rev* (2020) 57:100962. Epub 2019/09/11. doi: 10.1016/j.arr.2019.100962. PubMed PMID: 31505259.

47. Ihle-Hansen H, Thommessen B, Fagerland MW, Oksengard AR, Wyller TB, Engedal K, et al. Multifactorial vascular risk factor intervention to prevent cognitive impairment after stroke and TIA: a 12-month randomized controlled trial. *Int J Stroke* (2014) 9(7):932-8. Epub 2012/12/05. doi: 10.1111/j.1747-4949.2012.00928.x. PubMed PMID: 23205666.

48. Matz K, Teuschl Y, Firlinger B, Dachenhausen A, Keindl M, Seyfang L, et al. Multidomain Lifestyle Interventions for the Prevention of Cognitive Decline After Ischemic Stroke: Randomized Trial. *Stroke* (2015) 46(10):2874-80. Epub 2015/09/17. doi: 10.1161/STROKEAHA.115.009992. PubMed PMID: 26374482.

49. Merriman NA, Sexton E, McCabe G, Walsh ME, Rohde D, Gorman A, et al. Addressing cognitive impairment following stroke: systematic review and meta-analysis of non-randomised controlled studies of psychological interventions. *BMJ Open* (2019) 9(2):e024429. Epub 2019/03/02. doi: 10.1136/bmjopen-2018-024429. PubMed PMID: 30819706.

50. O'Brien JT, Thomas A. Vascular dementia. *Lancet* (2015) 386(10004):1698-706. Epub 2015/11/26. doi: 10.1016/S0140-6736(15)00463-8. PubMed PMID: 26595643.

51. Desmond DW. The neuropsychology of vascular cognitive impairment: is there a specific cognitive deficit? *J Neurol Sci* (2004) 226(1-2):3-7. Epub 2004/11/13. doi: 10.1016/j.jns.2004.09.002. PubMed PMID: 15537510.

52. Stephens S, Kenny RA, Rowan E, Allan L, Kalaria RN, Bradbury M, et al. Neuropsychological characteristics of mild vascular cognitive impairment and dementia after stroke. *Int J Geriatr Psychiatry* (2004) 19(11):1053-7. Epub 2004/10/14. doi: 10.1002/gps.1209. PubMed PMID: 15481073.

53. O'Brien JT. Vascular cognitive impairment. *Am J Geriatr Psychiatry* (2006) 14(9):724-33. Epub 2006/09/01. doi: 10.1097/01.JGP.0000231780.44684.7e. PubMed PMID: 16943169.

54. Gaebel W, Reed GM, Jakob R. Neurocognitive disorders in ICD-11: a new proposal and its outcome. *World Psychiatry* (2019) 18(2):232-3. Epub 2019/05/07. doi: 10.1002/wps.20634. PubMed PMID: 31059606.

55. Hachinski VC, Iliff LD, Zilhka E, Du Boulay GH, McAllister VL, Marshall J, et al. Cerebral blood flow in dementia. *Arch Neurol* (1975) 32(9):632-7. Epub 1975/09/01. doi: 10.1001/archneur.1975.00490510088009. PubMed PMID: 1164215.

56. Roman GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology* (1993) 43(2):250-60. Epub 1993/02/01. doi: 10.1212/wnl.43.2.250. PubMed PMID: 8094895.

57. van Straaten EC, Scheltens P, Knol DL, van Buchem MA, van Dijk EJ, Hofman PA, et al. Operational definitions for the NINDS-AIREN criteria for vascular dementia: an interobserver study. *Stroke* (2003) 34(8):1907-12. Epub 2003/07/12. doi: 10.1161/01.STR.0000083050.44441.10. PubMed PMID: 12855825.

58. McAleese KE, Alafuzoff I, Charidimou A, De Reuck J, Grinberg LT, Hainsworth AH, et al. Post-mortem assessment in vascular dementia: advances and aspirations. *BMC Med* (2016) 14(1):129. Epub 2016/09/08. doi: 10.1186/s12916-016-0676-5. PubMed PMID: 27600683.

59. Wiederkehr S, Simard M, Fortin C, van Reekum R. Comparability of the clinical diagnostic criteria for vascular dementia: a critical review. Part I. *J Neuropsychiatry Clin Neurosci* (2008) 20(2):150-61. Epub 2008/05/03. doi: 10.1176/appi.neuropsych.20.2.150. PubMed PMID: 18451186.

60. Wiederkehr S, Simard M, Fortin C, van Reekum R. Validity of the clinical diagnostic criteria for vascular dementia: A critical review. Part II. *Journal of Neuropsychiatry and Clinical Neurosciences* (2008) 20(2):162-77. doi: DOI 10.1176/appi.neuropsych.20.2.162. PubMed PMID: WOS:000257991900004.

61. Sachdev PS, Lipnicki DM, Crawford JD, Brodaty H. The Vascular Behavioral and Cognitive Disorders criteria for vascular cognitive disorders: a validation study. *European Journal of Neurology* (2019) 26(9):1161-7. doi: 10.1111/ene.13960. PubMed PMID: WOS:000478748700006.

62. Quinn TJ, Elliott E, Langhorne P. Cognitive and Mood Assessment Tools for Use in Stroke. *Stroke* (2018) 49(2):483-90. doi: 10.1161/Strokeaha.117.016994. PubMed PMID: WOS:000422928000046.

63. Cordell CB, Borson S, Boustani M, Chodosh J, Reuben D, Verghese J, et al. Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during the Medicare Annual Wellness Visit in a primary care setting. *Alzheimer's & dementia : the journal of the Alzheimer's Association* (2013) 9(2):141-50. Epub 2012/12/26. doi: 10.1016/j.jalz.2012.09.011. PubMed PMID: 23265826.

64. National Institute for Health and Care Excellence (NICE). Delirium: prevention, diagnosis and management (2010) [updated March 2019; cited 2021 March 5]. Available from: https://www.nice.org.uk/guidance/cg103/chapter/1-Guidance#treating-delirium.

65. Mulkey MA, Olson DM, Hardin SR. Top four evidence-based nursing interventions for delirium. *Medsurg Nursing* (2019) 28(6):357-62.

66. 4AT rapid clinical test for delirium [cited 2019 October 15]. Available from: https://www.the4at.com/.

67. Bellelli G, Morandi A, Davis DHJ, Mazzola P, Turco R, Gentile S, et al. Validation of the 4AT, a new instrument for rapid delirium screening: a study in 234 hospitalised older people. *Age Ageing* (2014) 43(4):496-502. doi: 10.1093/ageing/afu021.

68. Ely EW, Margolin R, Francis J, May L, Truman B, Dittus R, et al. Evaluation of delirium in critically ill patients: validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). *Crit Care Med* (2001) 29(7):1370-9. Epub 2001/07/11. doi: 10.1097/00003246-200107000-00012. PubMed PMID: 11445689.

69. Mansutti I, Saiani L, Palese A. Detecting delirium in patients with acute stroke: a systematic review of test accuracy. *Bmc Neurol* (2019) 19(1):310. Epub 2019/12/04. doi: 10.1186/s12883-019-1547-4. PubMed PMID: 31791260.

70. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet* (1974) 2(7872):81-4. Epub 1974/07/13. doi: 10.1016/s0140-6736(74)91639-0. PubMed PMID: 4136544.

71. Brott T, Adams HP, Jr., Olinger CP, Marler JR, Barsan WG, Biller J, et al. Measurements of acute cerebral infarction: a clinical examination scale. *Stroke* (1989) 20(7):864-70. Epub 1989/07/01. doi: 10.1161/01.str.20.7.864. PubMed PMID: 2749846.

72. Stroke Scales and Related Information: National Institute of Health, National Institute of Neurological Disorders and Stroke. [updated March 29, 2016; cited 2019 September 20]. Available from: https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Preventing-Stroke/Stroke-Scales-and-Related-Information.

73. Cote R, Hachinski VC, Shurvell BL, Norris JW, Wolfson C. The Canadian Neurological Scale: a preliminary study in acute stroke. *Stroke* (1986) 17(4):731-7. Epub 1986/07/01. doi: 10.1161/01.str.17.4.731. PubMed PMID: 3738958.

74. Elliott E, Drozdowska BA, Taylor-Rowan M, Shaw RC, Cuthbertson G, Quinn TJ. Who Is Classified as Untestable on Brief Cognitive Screens in an Acute Stroke Setting? *Diagnostics (Basel)* (2019) 9(3). Epub 2019/08/17. doi: 10.3390/diagnostics9030095. PubMed PMID: 31416176.

75. Stolwyk RJ, O'Neill MH, McKay AJD, Wong DK. Are Cognitive Screening Tools Sensitive and Specific Enough for Use After Stroke? A Systematic Literature Review. *Stroke* (2014) 45(10):3129-+. doi: 10.1161/Strokeaha.114.004232. PubMed PMID: WOS:000342794700060.

76. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* (1975) 12(3):189-98. Epub 1975/11/01. doi: 10.1016/0022-3956(75)90026-6. PubMed PMID: 1202204.

77. Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* (2005) 53(4):695-9. Epub 2005/04/09. doi: 10.1111/j.1532-5415.2005.53221.x. PubMed PMID: 15817019.

78. Mitchell AJ. A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *J Psychiatr Res* (2009) 43(4):411-31. Epub 2008/06/27. doi: 10.1016/j.jpsychires.2008.04.014. PubMed PMID: 18579155.

79. Lees R, Selvarajah J, Fenton C, Pendlebury ST, Langhorne P, Stott DJ, et al. Test accuracy of cognitive screening tests for diagnosis of dementia and multidomain cognitive impairment in stroke. *Stroke* (2014) 45(10):3008-18. Epub 2014/09/06. doi: 10.1161/STROKEAHA.114.005842. PubMed PMID: 25190446.

80. Chu K. An introduction to sensitivity, specificity, predictive values and likelihood ratios. *Emergency Medicine* (1999) 11(3):175-81.

81. Cullen B, O'Neill B, Evans JJ, Coen RF, Lawlor BA. A review of screening tests for cognitive impairment. *J Neurol Neurosurg Psychiatry* (2007) 78(8):790-9. Epub 2006/12/21. doi: 10.1136/jnnp.2006.095414. PubMed PMID: 17178826.

82. Swift A, Heale R, Twycross A. What are sensitivity and specificity? *Evid Based Nurs* (2020) 23(1):2-4. Epub 2019/11/14. doi: 10.1136/ebnurs-2019-103225. PubMed PMID: 31719126.

83. Roebuck-Spencer TM, Glen T, Puente AE, Denney RL, Ruff RM, Hostetter G, et al. Cognitive Screening Tests Versus Comprehensive Neuropsychological Test Batteries: A National Academy of Neuropsychology Education Paperdagger. *Arch Clin Neuropsychol* (2017) 32(4):491-8. Epub 2017/03/24. doi: 10.1093/arclin/acx021. PubMed PMID: 28334244.

84. National Institute for Health and Care Excellence (NICE). Stroke rehabilitation in adults.: NICE (2013) [cited 2021 March 3]. Available from: https://www.nice.org.uk/guidance/cg162/chapter/1-recommendations#cognitive-functioning-2.

85. Hachinski V, Iadecola C, Petersen RC, Breteler MM, Nyenhuis DL, Black SE, et al. National Institute of Neurological Disorders and Stroke-Canadian Stroke Network vascular cognitive impairment harmonization standards. *Stroke* (2006) 37(9):2220-41. Epub 2006/08/19. doi: 10.1161/01.STR.0000237236.88823.47. PubMed PMID: 16917086.

86. Ahmed N, Steiner T, Caso V, Wahlgren N, participants E-Ks. Recommendations from the ESO-Karolinska Stroke Update Conference, Stockholm 13-15 November 2016. *Eur Stroke J* (2017) 2(2):95-102. Epub 2018/06/15. doi: 10.1177/2396987317699144. PubMed PMID: 29900406.

87. Lees R, Fearon P, Harrison JK, Broomfield NM, Quinn TJ. Cognitive and mood assessment in stroke research: focused review of contemporary studies. *Stroke* (2012) 43(6):1678-80. Epub 2012/04/27. doi: 10.1161/STROKEAHA.112.653303. PubMed PMID: 22535271.

88. Saa JP, Tse T, Baum C, Cumming T, Josman N, Rose M, et al. Longitudinal evaluation of cognition after stroke - A systematic scoping review. *PLoS One* (2019) 14(8):e0221735. Epub 2019/08/30. doi: 10.1371/journal.pone.0221735. PubMed PMID: 31465492.

89. Demeyere N, Riddoch MJ, Slavkova ED, Bickerton WL, Humphreys GW. The Oxford Cognitive Screen (OCS): validation of a stroke-specific short cognitive screening tool. *Psychol Assess* (2015) 27(3):883-94. Epub 2015/03/03. doi: 10.1037/pas0000082. PubMed PMID: 25730165.

90. Crivelli D, Angelillo MT, Grippa E, Colucci A, Nardulli R, Balconi M. When is a novel psychometric measure needed? A preliminary analysis regarding the Cognitive Assessment for Stroke Patients (CASP) battery compared with MMSE and MoCA. *Appl Neuropsych-Adul* (2018) 25(5):410-6. doi: 10.1080/23279095.2017.1320556. PubMed PMID: WOS:000436574500004.

91. Qian L, Ding L, Cheng L, Zhu X, Zhao H, Jin J, et al. Early biomarkers for post-stroke cognitive impairment. *J Neurol* (2012) 259(10):2111-8. Epub 2012/03/13. doi: 10.1007/s00415-012-6465-y. PubMed PMID: 22407276.

92. Liu M, Zhou K, Li H, Dong X, Tan G, Chai Y, et al. Potential of serum metabolites for diagnosing post-stroke cognitive impairment. *Mol Biosyst* (2015) 11(12):3287-96. Epub 2015/10/23. doi: 10.1039/c5mb00470e. PubMed PMID: 26490688.

93. Chen YC, Chou WH, Fang CP, Liu TH, Tsou HH, Wang Y, et al. Serum Level and Activity of Butylcholinesterase: A Biomarker for Post-Stroke Dementia. *J Clin Med* (2019) 8(11). Epub 2019/10/28. doi: 10.3390/jcm8111778. PubMed PMID: 31653081.

94. Grobbee DE, Hoes AW. *Clinical epidemiology: principles, methods, and applications for clinical research*. Burlington, MA: Jones & Bartlett Publishers (2014).

95. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann Intern Med* (2019) 170(1):51-8. Epub 2019/01/01. doi: 10.7326/M18-1376. PubMed PMID: 30596875.

96. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ* (2013) 346:e5595. Epub 2013/02/07. doi: 10.1136/bmj.e5595. PubMed PMID: 23386360.

97. Kent P, Cancelliere C, Boyle E, Cassidy JD, Kongsted A. A conceptual framework for prognostic research. *BMC Med Res Methodol* (2020) 20(1):172. Epub 2020/07/01. doi: 10.1186/s12874-020-01050-7. PubMed PMID: 32600262.

98. University College London. Prognosis research partnership: UCL [cited 2021 March 6]. Available from: https://www.ucl.ac.uk/health-informatics/research/impact-research/prognosis-research-partnership.

99. Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med* (2013) 10(2):e1001380. Epub 2013/02/09. doi: 10.1371/journal.pmed.1001380. PubMed PMID: 23393429.

100. Hughes TF, Ganguli M. Modifiable Midlife Risk Factors for Late-Life Cognitive Impairment and Dementia. *Curr Psychiatry Rev* (2009) 5(2):73-92. Epub 2009/12/01. doi: 10.2174/157340009788167347. PubMed PMID: 19946443.

101. Regitz-Zagrosek V, Oertelt-Prigione S, Seeland U, Hetzer R. Sex and gender differences in myocardial hypertrophy and heart failure. *Circ J* (2010) 74(7):1265-73. Epub 2010/06/19. doi: 10.1253/circj.cj-10-0196. PubMed PMID: 20558892.

102. World Health Organization (WHO). Global Health Observatory (GHO) data. Female life expectancy.: WHO [cited 2021 March 2]. Available from: https://www.who.int/gho/women_and_health/mortality/life_expectancy_text/en/.

103. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ* (2009) 338:b375. Epub 2009/02/25. doi: 10.1136/bmj.b375. PubMed PMID: 19237405.

104. Herbert RD. Cohort studies of aetiology and prognosis: they're different. *J Physiother* (2014) 60(4):241-4. Epub 2014/12/03. doi: 10.1016/j.jphys.2014.07.005. PubMed PMID: 25443537.

105. Hingorani AD, Windt DA, Riley RD, Abrams K, Moons KG, Steyerberg EW, et al. Prognosis research strategy (PROGRESS) 4: stratified medicine research. *BMJ* (2013) 346:e5793. Epub 2013/02/07. doi: 10.1136/bmj.e5793. PubMed PMID: 23386361.

106. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med* (2013) 10(2):e1001381. Epub 2013/02/09. doi: 10.1371/journal.pmed.1001381. PubMed PMID: 23393430.

107. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: validating a prognostic model. *Bmj-Brit Med J* (2009) 338. doi: ARTN b605; 10.1136/bmj.b605. PubMed PMID: WOS:000266502400003.

108. Ntaios G, Faouzi M, Ferrari J, Lang W, Vemmos K, Michel P. An integer-based score to predict functional outcome in acute ischemic stroke: the ASTRAL score. *Neurology* (2012) 78(24):1916-22. doi: 10.1212/WNL.0b013e318259e221.

109. Steyerberg EW. *Clinical prediction models: a practical approach to development, validation, and updating*: Springer Science & Business Media (2009).

110. Pencina MJ, D'Agostino RB, Sr. Evaluating Discrimination of Risk Prediction Models: The C Statistic. *JAMA* (2015) 314(10):1063-4. Epub 2015/09/09. doi: 10.1001/jama.2015.11082. PubMed PMID: 26348755.

111. Crowson CS, Atkinson EJ, Therneau TM. Assessing calibration of prognostic risk scores. *Stat Methods Med Res* (2016) 25(4):1692-706. Epub 2013/08/03. doi: 10.1177/0962280213497434. PubMed PMID: 23907781.

112. Carter JV, Pan JM, Rai SN, Galandiuk S. ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. *Surgery* (2016) 159(6):1638-45. doi: 10.1016/j.surg.2015.12.029. PubMed PMID: WOS:000375890400017.

113. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* (2010) 21(1):128-38. Epub 2009/12/17. doi: 10.1097/EDE.0b013e3181c30fb2. PubMed PMID: 20010215.

114. Hosmer DW, Lemeshow S. *Applied logistic regression*. 2nd ed. New York: Wiley (2000). xii, 373 p. p.

115. Minne L, Ludikhuize J, de Rooij SEJA, Abu-Hanna A. Characterizing Predictive Models of Mortality for Older Adults and Their Validation for Use in Clinical Practice. *J Am Geriatr Soc* (2011) 59(6):1110-5. doi: 10.1111/j.1532-5415.2011.03411.x. PubMed PMID: WOS:000291978300020.

116. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. *Ann Intern Med* (2019) 170(1):W1-W33. Epub 2019/01/01. doi: 10.7326/M18-1377. PubMed PMID: 30596876.

117. Steyerberg EW, Harrell FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* (2001) 54(8):774-81. doi: Doi 10.1016/S0895-4356(01)00341-9. PubMed PMID: WOS:000170019900003.

118. Efron B, Tibshirani R. *An introduction to the bootstrap*. New York: Chapman & Hall (1993).

119. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* (2000) 19(4):453-73. doi: Doi 10.1002/(Sici)1097-0258(20000229)19:4<453::Aid-Sim350>3.3.Co;2-X. PubMed PMID: WOS:000085613900005.

120. Ramspek CL, Jager KJ, Dekker FW, Zoccali C, van Diepen M. External validation of prognostic models: what, why, how, when and where? *Clinical Kidney Journal* (2021) 14(1):49-58. doi: 10.1093/ckj/sfaa188.

121. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* (1999) 130(6):515-24. Epub 1999/03/13. doi: 10.7326/0003-4819-130-6-199903160-00016. PubMed PMID: 10075620.

122. Austin PC, van Klaveren D, Vergouwe Y, Nieboer D, Lee DS, Steyerberg EW. Geographic and temporal validity of prediction models: different approaches were useful to examine model performance. *J Clin Epidemiol* (2016) 79:76-85. doi: 10.1016/j.jclinepi.2016.05.007. PubMed PMID: WOS:000389967700013.

123. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. *Stat Med* (2004) 23(16):2567-86. Epub 2004/08/03. doi: 10.1002/sim.1844. PubMed PMID: 15287085.

124. Sim J, Teece L, Dennis MS, Roffe C, Team SOSS. Validation and Recalibration of Two Multivariable Prognostic Models for Survival and Independence in Acute Stroke. *PLoS One* (2016) 11(5):e0153527. Epub 2016/05/27. doi: 10.1371/journal.pone.0153527. PubMed PMID: 27227988.

125. Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic medicine* (2004) 66(3):411-21.

126. Frisoni GB, Scheltens P, Galluzzi S, Nobili FM, Fox NC, Robert PH, et al. Neuroimaging tools to rate regional atrophy, subcortical cerebrovascular disease, and regional cerebral blood flow and metabolism: consensus paper of the EADC. *J Neurol Neurosurg Psychiatry* (2003) 74(10):1371-81. Epub 2003/10/23. doi: 10.1136/jnnp.74.10.1371. PubMed PMID: 14570828.

127. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ* (2009) 338:b606. Epub 2009/06/09. doi: 10.1136/bmj.b606. PubMed PMID: 19502216.

128. Reilly BM, Evans AT. Translating clinical research into clinical practice: Impact of using prediction rules to make decisions. *Annals of Internal Medicine* (2006) 144(3):201-9. doi: 10.7326/0003-4819-144-3-200602070-00009. PubMed PMID: WOS:000235074800007.

129. Genuis SJ, Genuis SK. Resisting cookbook medicine. *Bmj* (2004) 329(7458):179.

130. How to calculate risk: BMJ [cited 2021 March 8]. Available from: https://bestpractice.bmj.com/info/toolkit/learn-ebm/how-to-calculate-risk/.

131. National Institute for Health and Care Excellence (NICE). Atrial fibrillation: managament: NICE (2014) [updated August 1 2014; cited 2021 March 9]. Available from: https://www.nice.org.uk/guidance/cg180/chapter/1-Recommendations#management-for-people-presenting-acutely-with-atrial-fibrillation.

132. Seelig J, Pisters R, Hemels ME, Huisman MV, ten Cate H, Alings M. When to withhold oral anticoagulation in atrial fibrillation - an overview of frequent clinical discussion topics. *Vascular Health and Risk Management* (2019) 15:399-408. doi: 10.2147/Vhrm.S187656. PubMed PMID: WOS:000486212500001.

133. Fava A, Petri M. Systemic lupus erythematosus: Diagnosis and clinical management. *J Autoimmun* (2019) 96:1-13. Epub 2018/11/19. doi: 10.1016/j.jaut.2018.11.001. PubMed PMID: 30448290.

134. Belle DJ, Singh H. Genetic factors in drug metabolism. *Am Fam Physician* (2008) 77(11):1553-60. Epub 2008/06/28. PubMed PMID: 18581835.

135. Hudis CA. Trastuzumab - mechanism of action and use in clinical practice. *N Engl J Med* (2007) 357(1):39-51. Epub 2007/07/06. doi: 10.1056/NEJMra043186. PubMed PMID: 17611206.

136. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *Arch Pathol Lab Med* (2007) 131(1):18-43. Epub 2007/01/01. doi: 10.1043/1543-2165(2007)131[18:ASOCCO]2.0.CO;2. PubMed PMID: 19548375.

137. Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. Refining Clinical Risk Stratification for Predicting Stroke and Thromboembolism in Atrial Fibrillation Using a Novel Risk Factor-Based Approach The Euro Heart Survey on Atrial Fibrillation. *Chest* (2010) 137(2):263-72. doi: 10.1378/chest.09-1584. PubMed PMID: WOS:000274612600006.

138. Medicines Management and Therapeutics Committee. *Atrial fibrillation and anticoagulation management*. In: Medicines Information Leaflet (2019) 8(5); Publisher: Oxford University Hospitals NHS Foundation Trust.

139. Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJ, Lip GY. A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey. *Chest* (2010) 138(5):1093-100. Epub 2010/03/20. doi: 10.1378/chest.10-0134. PubMed PMID: 20299623.

140. De Backer O, Arnous S, Ihlemann N, Vejlstrup N, Jorgensen E, Pehrson S, et al. Percutaneous left atrial appendage occlusion for stroke prevention in atrial fibrillation: an update. *Open Heart* (2014) 1(1):e000020. Epub 2014/10/22. doi: 10.1136/openhrt-2013-000020. PubMed PMID: 25332785.

141. Zhang X, Bi X. Post-Stroke Cognitive Impairment: A Review Focusing on Molecular Biomarkers. *J Mol Neurosci* (2020) 70(8):1244-54. Epub 2020/03/29. doi: 10.1007/s12031-020-01533-8. PubMed PMID: 32219663.

142. Casolla B, Caparros F, Cordonnier C, Bombois S, Henon H, Bordet R, et al. Biological and imaging predictors of cognitive impairment after stroke: a systematic review. *J Neurol* (2019) 266(11):2593-604. Epub 2018/10/24. doi: 10.1007/s00415-018-9089-z. PubMed PMID: 30350168.

143. Fong TG, Davis D, Growdon ME, Albuquerque A, Inouye SK. The interface between delirium and dementia in elderly adults. *Lancet Neurol* (2015) 14(8):823-32. doi: 10.1016/S1474-4422(15)00101-5. PubMed PMID: WOS:000357879200012.

144. Mansutti I, Saiani L, Palese A. Delirium in patients with ischaemic and haemorrhagic stroke: findings from a scoping review. *European Journal of Cardiovascular Nursing* (2019) 18(6):435-48. doi: 10.1177/1474515119846226. PubMed PMID: WOS:000477647400002.

145. Gottesman RF, Hillis AE. Predictors and assessment of cognitive dysfunction resulting from ischaemic stroke. *Lancet Neurol* (2010) 9(9):895-905. Epub 2010/08/21. doi: 10.1016/S1474-4422(10)70164-2. PubMed PMID: 20723846.

146. Danovska M, Peychinska D. Post-stroke cognitive impairment–phenomenology and prognostic factors. *Journal of IMAB–Annual Proceeding Scientific Papers* (2012) 18(3):290-7.

147. Carin-Levy G, Mead GE, Nicol K, Rush R, van Wijck F. Delirium in acute stroke: screening tools, incidence rates and predictors: a systematic review. *J Neurol* (2012) 259(8):1590-9. Epub 2012/01/12. doi: 10.1007/s00415-011-6383-4. PubMed PMID: 22234842.

148. OCEBM Levels of Evidence Working Group. *The Oxford 2011 Levels of Evidence*. In: Oxford Centre for Evidence-Based Medicine, editor. (2011).

149. Burns PB, Rohrich RJ, Chung KC. The Levels of Evidence and Their Role in Evidence-Based Medicine. *Plast Reconstr Surg* (2011) 128(1):305-10. doi: 10.1097/PRS.0b013e318219c171. PubMed PMID: WOS:000292499600068.

150. Tang EYH, Robinson L, Stephan BCM. Risk Prediction Models for Post-Stroke Dementia. *Geriatrics (Basel)* (2017) 2(3). Epub 2017/06/22. doi: 10.3390/geriatrics2030019. PubMed PMID: 31011029.

151. Drozdowska BA, McGill K, McKay M, Bartlam R, Langhorne P, Quinn TJ. Prognostic rules for predicting cognitive syndromes following stroke: A systematic review. *Eur Stroke J* (2021) 6(1):18-27. Epub 2021/04/06. doi: 10.1177/2396987321997045. PubMed PMID: 33817331.

152. Hou XH, Feng L, Zhang C, Cao XP, Tan L, Yu JT. Models for predicting risk of dementia: a systematic review. *J Neurol Neurosurg Psychiatry* (2019) 90(4):373-9. Epub 2018/06/30. doi: 10.1136/jnnp-2018-318212. PubMed PMID: 29954871.

153. Lindroth H, Bratzke L, Purvis S, Brown R, Coburn M, Mrkobrada M, et al. Systematic review of prediction models for delirium in the older adult inpatient. *BMJ Open* (2018) 8(4):e019223. Epub 2018/05/01. doi: 10.1136/bmjopen-2017-019223. PubMed PMID: 29705752.

154. Tang EYH, Price CI, Robinson L, Exley C, Desmond DW, Kohler S, et al. Assessing the Predictive Validity of Simple Dementia Risk Models in Harmonized Stroke Cohorts. *Stroke* (2020) 51(7):2095-102. Epub 2020/06/23. doi: 10.1161/STROKEAHA.120.027473. PubMed PMID: 32568644.

155. Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med* (2014) 11(10):e1001744. Epub 2014/10/15. doi: 10.1371/journal.pmed.1001744. PubMed PMID: 25314315.

156. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* (2009) 151(4):264-9, W64. Epub 2009/07/23. doi: 10.7326/0003-4819-151-4-200908180-00135. PubMed PMID: 19622511.

157. Cochrane Stroke. Search Methods for the Cochrane Stroke Group Specialised Register. Resources for authors and editors of reviews. [cited 2019 October 7]. Available from: https://apps.ccbs.ed.ac.uk/csrg/entity/searchmethods.pdf.

158. Geersing GJ, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons KG. Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. *PLoS One* (2012) 7(2):e32844. Epub 2012/03/07. doi: 10.1371/journal.pone.0032844. PubMed PMID: 22393453.

159. Flodgren GM, Berg RC. Primary and secondary prevention interventions for cognitive decline and dementia. Oslo: (2016).

160. National Institute for Health and Care Excellence (NICE). Appendix D: Review search strategies. Dementia: assessment, management and support for people living with dementia and their carers. (2018) [cited 2019 October 7]. Available from: https://www.nice.org.uk/guidance/ng97/documents/search-strategies.

161. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev* (2016) 5(1):210. Epub 2016/12/07. doi: 10.1186/s13643-016-0384-4. PubMed PMID: 27919275.

162. Carter JV, Pan J, Rai SN, Galandiuk S. ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. *Surgery* (2016) 159(6):1638-45. Epub 2016/03/11. doi: 10.1016/j.surg.2015.12.029. PubMed PMID: 26962006.

163. Drozdowska BA, Singh S, Quinn TJ. Thinking About the Future: A Review of Prognostic Scales Used in Acute Stroke. *Frontiers in Neurology* (2019) 10. doi: ARTN 274; 10.3389/fneur.2019.00274. PubMed PMID: WOS:000461859800001.

164. Chander RJ, Lam BYK, Lin X, Ng AYT, Wong APL, Mok VCT, et al. Development and validation of a risk score (CHANGE) for cognitive impairment after ischemic stroke. *Sci Rep* (2017) 7(1):12441. Epub 2017/10/01. doi: 10.1038/s41598-017-12755-z. PubMed PMID: 28963553.

165. Ding MY, Xu Y, Wang YZ, Li PX, Mao YT, Yu JT, et al. Predictors of Cognitive Impairment After Stroke: A Prospective Stroke Cohort Study. *J Alzheimers Dis* (2019) 71(4):1139-51. Epub 2019/09/17. doi: 10.3233/JAD-190382. PubMed PMID: 31524163.

166. Gong K, Zhao L, Guo J, Wang Z. A nomogram to predict cognitive impairment after supratentorial spontaneous intracranial hematoma in adult patients: A retrospective cohort study. *Medicine (Baltimore)* (2019) 98(42):e17626. Epub 2019/10/19. doi: 10.1097/MD.0000000000017626. PubMed PMID: 31626144.

167. Kandiah N, Chander RJ, Lin X, Ng A, Poh YY, Cheong CY, et al. Cognitive Impairment after Mild Stroke: Development and Validation of the SIGNAL2 Risk Score. *J Alzheimers Dis* (2016) 49(4):1169-77. Epub 2015/11/26. doi: 10.3233/JAD-150736. PubMed PMID: 26599056.

168. Lin JH, Lin RT, Tai CT, Hsieh CL, Hsiao SF, Liu CK. Prediction of poststroke dementia. *Neurology* (2003) 61(3):343-8. Epub 2003/08/13. doi: 10.1212/01.wnl.0000078891.27052.10. PubMed PMID: 12913195.

169. Munsch F, Sagnier S, Asselineau J, Bigourdan A, Guttmann CR, Debruxelles S, et al. Stroke Location Is an Independent Predictor of Cognitive Outcome. *Stroke* (2016) 47(1):66-73. doi: 10.1161/Strokeaha.115.011242. PubMed PMID: WOS:000367136500011.

170. Salihovic D, Smajlovic D, Mijajlovic M, Zoletic E, Ibrahimagic OC. Cognitive syndromes after the first stroke. *Neurological Sciences* (2018) 39(8):1445-51. doi: 10.1007/s10072-018-3447-6. PubMed PMID: WOS:000439465800016.

171. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* (1988) 19(5):604-7. Epub 1988/05/01. doi: 10.1161/01.str.19.5.604. PubMed PMID: 3363593.

172. Kostalova M, Bednarik J, Mitasova A, Dusek L, Michalcakova R, Kerkovsky M, et al. Towards a predictive model for post-stroke delirium. *Brain Injury* (2012) 26(7-8):962-71. doi: 10.3109/02699052.2012.660510. PubMed PMID: WOS:000305135100007.

173. Kotfis K, Bott-Olejnik M, Szylinska A, Rotter I. Could Neutrophil-to-Lymphocyte Ratio (NLR) Serve as a Potential Marker for Delirium Prediction in Patients with Acute Ischemic Stroke? A Prospective Observational Study. *Journal of Clinical Medicine* (2019) 8(7). doi: ARTN 1075; 10.3390/jcm8071075. PubMed PMID: WOS:000479003300157.

174. Oldenbeuving AW, de Kort PLM, van der Sluijs JFV, Kappelle LJ, Roks G. An early prediction of delirium in the acute phase after stroke. *J Neurol Neurosur Ps* (2014) 85(4):431-4. doi: 10.1136/jnnp-2013-304920. PubMed PMID: WOS:000332308500016.

175. Tang E, Exley C, Price C, Stephan B, Robinson L. The views of public and clinician stakeholders on risk assessment tools for post-stroke dementia: a qualitative study. *BMJ Open* (2019) 9(3):e025586. Epub 2019/03/29. doi: 10.1136/bmjopen-2018-025586. PubMed PMID: 30918033.

176. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* (2015) 44(3):827-36. Epub 2015/06/08. doi: 10.1093/ije/dyv098. PubMed PMID: 26050254.

177. Riley RD, Ensor J, Snell KIE, Debray TPA, Altman DG, Moons KGM, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. *Bmj-Brit Med J* (2016) 353. doi: ARTN i3140; 10.1136/bmj.i3140. PubMed PMID: WOS:000378723500006.

178. Brainin M, Tuomilehto J, Heiss WD, Bornstein NM, Bath PMW, Teuschl Y, et al. Post-stroke cognitive decline: an update and perspectives for clinical research. *European Journal of Neurology* (2015) 22(2):229. doi: 10.1111/ene.12626. PubMed PMID: WOS:000347701300004.

179. Sun GW, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol* (1996) 49(8):907-16. doi: Doi 10.1016/0895-4356(96)00025-X. PubMed PMID: WOS:A1996VB72900014.

180. Hayden JA, Cote P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann Intern Med* (2006) 144(6):427-37. Epub 2006/03/22. doi: 10.7326/0003-4819-144-6-200603210-00010. PubMed PMID: 16549855.

181. Young RJ. The effect of regular exercise on cognitive functioning and personality. *Br J Sports Med* (1979) 13(3):110-7. Epub 1979/09/01. doi: 10.1136/bjsm.13.3.110. PubMed PMID: 486882.

182. Monk A. Aging, Loneliness, and Communications. *Am Behav Sci* (1988) 31(5):532-63. doi: Doi 10.1177/000276488031005004. PubMed PMID: WOS:A1988N948700003.

183. Gajewski PD, Falkenstein M. Physical activity and neurocognitive functioning in aging - a condensed updated review. *Eur Rev Aging Phys A* (2016) 13. doi: ARTN 1; 10.1186/s11556-016-0161-3. PubMed PMID: WOS:000369694000001.

184. Erickson KI, Hillman C, Sullman CM, Ballard RM, Bloodgood B, Conroy DE, et al. Physical Activity, Cognition, and Brain Outcomes: A Review of the 2018 Physical Activity Guidelines. *Medicine and Science in Sports and Exercise* (2019) 51(6):1242-51. doi: 10.1249/Mss.0000000000001936. PubMed PMID: WOS:000468480400019.

185. Penninkilampi R, Casey AN, Singh MF, Brodaty H. The Association between Social Engagement, Loneliness, and Risk of Dementia: A Systematic Review and Meta-Analysis. *J Alzheimers Dis* (2018) 66(4):1619-33. doi: 10.3233/Jad-180439. PubMed PMID: WOS:000457776400027.

186. Evans IEM, Martyr A, Collins R, Brayne C, Clare L. Social Isolation and Cognitive Function in Later Life: A Systematic Review and Meta-Analysis. *J Alzheimers Dis* (2019) 70:S119-S44. doi: 10.3233/Jad-180501. PubMed PMID: WOS:000481670800012.

187. Sommerlad A, Sabia S, Singh-Manoux A, Lewis G, Livingston G. Association of social contact with dementia and cognition: 28-year follow-up of the Whitehall II cohort study. *Plos Medicine* (2019) 16(8). doi: ARTN e1002862; 10.1371/journal.pmed.1002862. PubMed PMID: WOS:000489050500019.

188. Fratiglioni L, Paillard-Borg S, Winblad B. An active and socially integrated lifestyle in late life might protect against dementia. *Lancet Neurol* (2004) 3(6):343-53. doi: Doi 10.1016/S1474-4422(04)00767-7. PubMed PMID: WOS:000221663000016.

189. Dause TJ, Kirby ED. Aging gracefully: social engagement joins exercise and enrichment as a key lifestyle factor in resistance to age-related cognitive decline. *Neural Regeneration Research* (2019) 14(1):39-42. doi: 10.4103/1673-5374.243698. PubMed PMID: WOS:000451333500006.

190. Caspersen CJ, Powell KE, Christenson GM. Physical-Activity, Exercise, and Physical-Fitness - Definitions and Distinctions for Health-Related Research. *Public Health Rep* (1985) 100(2):126-31. PubMed PMID: WOS:A1985AFX9800005.

191. National Health Service (NHS). Exercise [updated 2019 8 Oct; cited 2020 20 Oct]. Available from: https://www.nhs.uk/live-well/exercise/.

192. Atkin AJ, Gorely T, Clemes SA, Yates T, Edwardson C, Brage S, et al. Methods of Measurement in epidemiology: Sedentary Behaviour. *International Journal of Epidemiology* (2012) 41(5):1460-71. doi: 10.1093/ije/dys118. PubMed PMID: WOS:000309922700032.

193. Hallgren M, Dunstan DW, Owen N. Passive Versus Mentally Active Sedentary Behaviors and Depression. *Exerc Sport Sci Rev* (2020) 48(1):20-7. doi: 10.1249/Jes.0000000000000211. PubMed PMID: WOS:000502755700005.

194. Bath PA, Deeg D. Social engagement and health outcomes among older people: introduction to a special section. *European Journal of Ageing* (2005) 2(1):24-30. doi: 10.1007/s10433-005-0019-4.

195. Hall A, Wellman B. Social networks and social support. (1985).

196. Berkman LF. The relationship of social networks and social support to morbidity and mortality. (1985).

197. Bowling A, Browne PD. Social Networks, Health, and Emotional Well-Being among the Oldest in London. *J Gerontol* (1991) 46(1):S20-S32. doi: DOI 10.1093/geronj/46.1.S20. PubMed PMID: WOS:A1991FJ10900014.

198. Everard KM, Lach HW, Fisher EB, Baum MC. Relationship of activity and social support to the functional health of older adults. *J Gerontol B-Psychol* (2000) 55(4):S208-S12. doi: DOI 10.1093/geronb/55.4.S208. PubMed PMID: WOS:000088045700008.

199. Cohen S, Hoberman HM. Positive Events and Social Supports as Buffers of Life Change Stress. *J Appl Soc Psychol* (1983) 13(2):99-125. doi: DOI 10.1111/j.1559-1816.1983.tb02325.x. PubMed PMID: WOS:A1983RB53800001.

200. Young FW, Glasgow N. Voluntary social participation and health. *Res Aging* (1998) 20(3):339-62. doi: Doi 10.1177/0164027598203004. PubMed PMID: WOS:000073237100004.

201. Gilmour H. Social participation and the health and well-being of Canadian seniors. *Health Rep* (2012) 23(4):23-32. PubMed PMID: WOS:000312760600003.

202. Kamiya Y, Whelan B, Timonen V, Kenny RA. The differential impact of subjective and objective aspects of social engagement on cardiovascular risk factors. *Bmc Geriatr* (2010) 10. doi: Artn 81; 10.1186/1471-2318-10-81. PubMed PMID: WOS:000208731600081.

203. Rook KS. Research on social support, loneliness, and social isolation: Toward an integration. *Review of Personality & Social Psychology* (1984).

204. Miles L. Physical activity and health. *Nutrition Bulletin* (2007) 32(4):314-63. doi: 10.1111/j.1467-3010.2007.00668.x.

205. Saunders TJ, McIsaac T, Douillette K, Gaulton N, Hunter S, Rhodes RE, et al. Sedentary behaviour and health in adults: an overview of systematic reviews. *Appl Physiol Nutr Metab* (2020) 45(10 (Suppl. 2)):S197-S217. Epub 2020/10/16. doi: 10.1139/apnm-2020-0272. PubMed PMID: 33054341.

206. Saunders TJ, Larouche R, Colley RC, Tremblay MS. Acute sedentary behaviour and markers of cardiometabolic risk: a systematic review of intervention studies. *J Nutr Metab* (2012) 2012:712435. Epub 2012/07/04. doi: 10.1155/2012/712435. PubMed PMID: 22754695.

207. Stamatakis E, de Rezende LFM, Rey-Lopez JP. Sedentary Behaviour and Cardiovascular Disease. *Springer Ser Epidemi* (2018):215-43. doi: 10.1007/978-3-319-61552-3_9. PubMed PMID: WOS:000465573600010.

208. Chan JSY, Liu G, Liang D, Deng K, Wu J, Yan JH. Special Issue - Therapeutic Benefits of Physical Activity for Mood: A Systematic Review on the Effects of Exercise Intensity, Duration, and Modality. *J Psychol* (2019) 153(1):102-25. Epub 2018/10/16. doi: 10.1080/00223980.2018.1470487. PubMed PMID: 30321106.

209. Berkman LF, Krishna A. Social Network Epidemiology. In: Berkman LF, Kawachi I, Glymour MM, editors. *Social Epidemiology, 2nd Edition*. New York: Oxford University Press (2014).

210. Uchino BN, Cacioppo JT, KiecoltGlaser JK. The relationship between social support and physiological processes: A review with emphasis on underlying mechanisms and implications for health. *Psychol Bull* (1996) 119(3):488-531. doi: Doi 10.1037/0033-2909.119.3.488. PubMed PMID: WOS:A1996UL59200008.

211. McEwen BS. The neurobiology of stress: from serendipity to clinical relevance. *Brain Res* (2000) 886(1-2):172-89. doi: Doi 10.1016/S0006-8993(00)02950-4. PubMed PMID: WOS:000166027200014.

212. Friedler B, Crapser J, McCullough L. One is the deadliest number: the detrimental effects of social isolation on cerebrovascular diseases and cognition. *Acta Neuropathol* (2015) 129(4):493-509. doi: 10.1007/s00401-014-1377-9. PubMed PMID: WOS:000351517200002.

213. Cohen S, Doyle WJ, Skoner DP, Rabin BS, Gwaltney JM. Social ties and susceptibility to the common cold. *Jama-J Am Med Assoc* (1997) 277(24):1940-4. PubMed PMID: WOS:A1997XF08700038.

214. Knox SS, Uvnas-Moberg K. Social isolation and cardiovascular disease: an atherosclerotic pathway? *Psychoneuroendocrinology* (1998) 23(8):877-90. PubMed PMID: 9924742.

215. Loucks EB, Berkman LF, Gruenewald TL, Seeman TE. Relation of social integration to inflammatory marker concentrations in men and women 70 to 79 years. *Am J Cardiol* (2006) 97(7):1010-6. doi: 10.1016/j.amjcard.2005.10.043. PubMed PMID: WOS:000236708700017.

216. Cohen S. Psychosocial stress, social networks and susceptibility to infection. *The link between religion and health: Psychoneuroimmunology and the faith factor* (2002):101-23.

217. Cacioppo JT, Hawkley LC. Social isolation and health, with an emphasis on underlying mechanisms. *Perspect Biol Med* (2003) 46(3):S39-S52. doi: DOI 10.1353/pbm.2003.0049. PubMed PMID: WOS:000184458500004.

218. Stuller KA, Jarrett B, DeVries AC. Stress and social isolation increase vulnerability to stroke. *Exp Neurol* (2012) 233(1):33-9. doi: 10.1016/j.expneurol.2011.01.016. PubMed PMID: WOS:000300123900005.

219. Akerlind I, Hornquist JO. Loneliness and Alcohol-Abuse - a Review of Evidences of an Interplay. *Social Science & Medicine* (1992) 34(4):405-14. doi: Doi 10.1016/0277-9536(92)90300-F. PubMed PMID: WOS:A1992HJ54000007.

220. Dyal SR, Valente TW. A Systematic Review of Loneliness and Smoking: Small Effects, Big Implications. *Subst Use Misuse* (2015) 50(13):1697-716. doi: 10.3109/10826084.2015.1027933. PubMed PMID: WOS:000365684100009.

221. Pritchard ME, Yalch KL. Relationships among loneliness, interpersonal dependency, and disordered eating in young adults. *Pers Indiv Differ* (2009) 46(3):341-6. doi: 10.1016/j.paid.2008.10.027. PubMed PMID: WOS:000263406900016.

222. Levine MP. Loneliness and Eating Disorders. *J Psychol* (2012) 146(1-2):243-57. doi: 10.1080/00223980.2011.606435. PubMed PMID: WOS:000302607900016.

223. Prang KH, Berecki-Gisolf J, Newnam S. Recovery from musculoskeletal injury: the role of social support following a transport accident. *Health Qual Life Outcomes* (2015) 13:97. Epub 2015/07/04. doi: 10.1186/s12955-015-0291-8. PubMed PMID: 26138816.

224. Wu JR, Frazier SK, Rayens MK, Lennie TA, Chung ML, Moser DK. Medication Adherence, Social Support, and Event-Free Survival in Patients With Heart Failure. *Health Psychol* (2013) 32(6):637-46. doi: 10.1037/a0028527. PubMed PMID: WOS:000319547300004.

225. Cardoso-Moreno MJ, Tomas-Aragones L. The influence of perceived family support on post surgery recovery. *Psychol Health Med* (2017) 22(1):121-8. Epub 2016/02/24. doi: 10.1080/13548506.2016.1153680. PubMed PMID: 26899740.

226. Wylde V, Kunutsor S, Lenguerrand E, Jackson J, Blom AW, Beswick AD. Association of social support with patient-reported outcomes after joint replacement: a systematic review and meta-analysis. *Lancet Rheumatol* (2019) 1(3):E174-E86. doi: 10.1016/S2665-9913(19)30050-5. PubMed PMID: WOS:000547827900016.

227. van Praag H, Kempermann G, Gage FH. Neural consequences of environmental enrichment. *Nat Rev Neurosci* (2000) 1(3):191-8. Epub 2001/03/22. doi: 10.1038/35044558. PubMed PMID: 11257907.

228. Volkers KM, Scherder EJ. Impoverished environment, cognition, aging and dementia. *Rev Neurosci* (2011) 22(3):259-66. Epub 2011/05/20. doi: 10.1515/RNS.2011.026. PubMed PMID: 21591910.

229. Rosenich E, Hordacre B, Paquet C, Koblar SA, Hillier SL. Cognitive Reserve as an Emerging Concept in Stroke Recovery. *Neurorehabilitation and Neural Repair* (2020) 34(3):187-99. doi: Artn 1545968320907071; 10.1177/1545968320907071. PubMed PMID: WOS:000515468100001.

230. McDonald MW, Hayward KS, Rosbergen ICM, Jeffers MS, Corbett D. Is Environmental Enrichment Ready for Clinical Application in Human Post-stroke Rehabilitation? *Frontiers in Behavioral Neuroscience* (2018) 12. doi: ARTN 135; 10.3389/fnbeh.2018.00135. PubMed PMID: WOS:000438309100001.

231. Zhao Z, Sabirzhanov B, Wu J, Faden AI, Stoica BA. Voluntary Exercise Preconditioning Activates Multiple Antiapoptotic Mechanisms and Improves Neurological Recovery after Experimental Traumatic Brain Injury. *J Neurotrauma* (2015) 32(17):1347-60. Epub 2014/11/25. doi: 10.1089/neu.2014.3739. PubMed PMID: 25419789.

232. Reinholdsson M, Palstam A, Sunnerhagen KS. Prestroke physical activity could influence acute stroke severity (part of PAPSIGOT). *Neurology* (2018) 91(16):E146-1467. doi: 10.1212/Wnl.0000000000006354. PubMed PMID: WOS:000452502900001.

233. Oberlin LE, Waiwood AM, Cumming TB, Marsland AL, Bernhardt J, Erickson KI. Effects of Physical Activity on Poststroke Cognitive Function A Meta-Analysis of Randomized Controlled Trials. *Stroke* (2017) 48(11):3093-+. doi: 10.1161/Strokeaha.117.017319. PubMed PMID: WOS:000413496900038.

234. Nicholson S, Sniehotta FF, van Wijck F, Greig CA, Johnston M, McMurdo MET, et al. A systematic review of perceived barriers and motivators to physical activity after stroke. *International Journal of Stroke* (2013) 8(5):357-64. doi: 10.1111/j.1747-4949.2012.00880.x. PubMed PMID: WOS:000320725200022.

235. Marzolini S, Fong K, Jagroop D, Neirinckx J, Liu J, Reyes R, et al. Eligibility, Enrollment, and Completion of Exercise-Based Cardiac Rehabilitation Following Stroke Rehabilitation: What Are the Barriers? *Phys Ther* (2020) 100(1):44-56. Epub 2019/10/08. doi: 10.1093/ptj/pzz149. PubMed PMID: 31588512.

236. Benatti FB, Ried-Larsen M. The Effects of Breaking up Prolonged Sitting Time: A Review of Experimental Studies. *Med Sci Sports Exerc* (2015) 47(10):2053-61. Epub 2015/09/18. doi: 10.1249/MSS.0000000000000654. PubMed PMID: 26378942.

237. Chrismas BCR, Taylor L, Cherif A, Sayegh S, Bailey DP. Breaking up prolonged sitting with moderate-intensity walking improves attention and executive function in Qatari females. *PLoS One* (2019) 14(7):e0219565. Epub 2019/07/13. doi: 10.1371/journal.pone.0219565. PubMed PMID: 31299061.

238. English C, Healy GN, Coates A, Lewis L, Olds T, Bernhardt J. Sitting and Activity Time in People With Stroke. *Physical Therapy* (2016) 96(2):193-201. doi: 10.2522/ptj.20140522. PubMed PMID: WOS:000369775800008.

239. English C, Janssen H, Crowfoot G, Bourne J, Callister R, Dunn A, et al. Frequent, short bouts of light-intensity exercises while standing decreases systolic blood pressure: Breaking Up Sitting Time after Stroke (BUST-Stroke) trial. *International Journal of Stroke* (2018) 13(9):932-40. doi: 10.1177/1747493018798535. PubMed PMID: WOS:000454363000009.

240. Falck RS, Davis JC, Liu-Ambrose T. What is the association between sedentary behaviour and cognitive function? A systematic review. *British Journal of Sports Medicine* (2017) 51(10):800-11. doi: 10.1136/bjsports-2015-095551. PubMed PMID: WOS:000400583100010.

241. Valtorta NK, Kanaan M, Gilbody S, Ronzi S, Hanratty B. Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. *Heart* (2016) 102(13):1009-16. doi: 10.1136/heartjnl-2015-308790. PubMed PMID: WOS:000378716400008.

242. Hommel M, Trabucco-Miguel S, Joray S, Naegele B, Gonnet N, Jaillard A, et al. Social dysfunctioning after mild to moderate first-ever stroke at vocational age. *J Neurol Neurosur Ps* (2009) 80(4):371-5. doi: 10.1136/jnnp.2008.157875. PubMed PMID: WOS:000264238800006.

243. Brady MC, Clark AM, Dickson S, Paton G, Barbour RS. The impact of stroke-related dysarthria on social participation and implications for rehabilitation. *Disabil Rehabil* (2011) 33(3):178-86. doi: 10.3109/09638288.2010.517897. PubMed PMID: WOS:000285553600002.

244. Mukherjee D, Levin RL, Heller W. The cognitive, emotional, and social sequelae of stroke: psychological and ethical concerns in post-stroke adaptation. *Top Stroke Rehabil* (2006) 13(4):26-35. Epub 2006/11/04. doi: 10.1310/tsr1304-26. PubMed PMID: 17082166.

245. Glymour MM, Weuve J, Fay ME, Glass T, Berkman LF. Social ties and cognitive recovery after stroke: Does social integration promote cognitive resilience? *Neuroepidemiology* (2008) 31(1):10-20. doi: 10.1159/000136646. PubMed PMID: WOS:000257115100002.

246. Glass TA, Matchar DB, Belyea M, Feussner JR. Impact of social support on outcome in first stroke. *Stroke* (1993) 24(1):64-70. Epub 1993/01/01. doi: 10.1161/01.str.24.1.64. PubMed PMID: 8418553.

247. Tsouna-Hadjis E, Vemmos KN, Zakopoulos N, Stamatelopoulos S. First-stroke recovery process: the role of family social support. *Arch Phys Med Rehabil* (2000) 81(7):881-7. Epub 2000/07/15. doi: 10.1053/apmr.2000.4435. PubMed PMID: 10895999.

248. Glass TA, Maddox GL. The quality and quantity of social support: stroke recovery as psycho-social transition. *Soc Sci Med* (1992) 34(11):1249-61. Epub 1992/06/01. doi: 10.1016/0277-9536(92)90317-j. PubMed PMID: 1641684.

249. Von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Plos Medicine* (2007) 4(10):1623-7. doi: ARTN e296; 10.1371/journal.pmed.0040296. PubMed PMID: WOS:000251113600014.

250. Drozdowska BA, Celis-Morales CA, Lyall DM, Quinn TJ. Social engagement after stroke–is it relevant to cognitive function? A cross-sectional analysis of UK Biobank data. *AMRC Open Research* (2019) 1(3):3.

251. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* (2015) 12(3):e1001779. doi: 10.1371/journal.pmed.1001779. PubMed PMID: 25826379.

252. Guyomard V, Metcalf AK, Naguib MF, Fulcher RA, Potter JF, Myint PK. Transient ischaemic attack, vascular risk factors and cognitive impairment: a case-controlled study. *Age Ageing* (2011) 40(5):641-4. Epub 2011/07/14. doi: 10.1093/ageing/afr079. PubMed PMID: 21749995.

253. Not just a funny turn. The real impact of TIA. Stroke Association, (2014).

254. Lyall DM, Celis-Morales CA, Anderson J, Gill JMR, Mackay DF, McIntosh AM, et al. Associations between single and multiple cardiometabolic diseases and cognitive abilities in 474 129 UK Biobank participants. *European Heart Journal* (2017) 38(8):577-83. doi: 10.1093/eurheartj/ehw528.

255. Cullen B, Nicholl BI, Mackay DF, Martin D, Ul-Haq Z, McIntosh A, et al. Cognitive function and lifetime features of depression and bipolar disorder in a large population sample: Cross-sectional study of 143,828 UK Biobank participants. *Eur Psychiat* (2015) 30(8):950-8. doi: 10.1016/j.eurpsy.2015.08.006. PubMed PMID: WOS:000365160100015.

256. Hagenaars SP, Harris SE, Davies G, Hill WD, Liewald DCM, Ritchie SJ, et al. Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N=112151) and 24 GWAS consortia. *Mol Psychiatr* (2016) 21(11):1624-32. doi: 10.1038/mp.2015.225. PubMed PMID: WOS:000388719600017.

257. Lyall DM, Cullen B, Allerhand M, Smith DJ, Mackay D, Evans J, et al. Cognitive Test Scores in UK Biobank: Data Reduction in 480,416 Participants and Longitudinal Stability in 20,346 Participants. *Plos One* (2016) 11(4). doi: ARTN e0154222; 10.1371/journal.pone.0154222. PubMed PMID: WOS:000374970600051.

258. Fawns-Ritchie C, Deary IJ. Reliability and validity of the UK Biobank cognitive tests. *Plos One* (2020) 15(4). doi: ARTN e0231627; 10.1371/journal.pone.0231627. PubMed PMID: WOS:000536019800023.

259. Townsend P, Phillimore P, Beattie A. *Health and deprivation: inequality and the North*: Routledge (1988).

260. Pendlebury ST. Stroke-related dementia: Rates, risk factors and implications for future research. *Maturitas* (2009) 64(3):165-71. doi: http://dx.doi.org/10.1016/j.maturitas.2009.09.010.

261. Kloppenborg RP, van den Berg E, Kappelle LJ, Biessels GJ. Diabetes and other vascular risk factors for dementia: Which factor matters most? A systematic review. *Eur J Pharmacol* (2008) 585(1):97-108. doi: 10.1016/j.ejphar.2008.02.049. PubMed PMID: WOS:000256149300012.

262. Skoog I. Risk-Factors for Vascular Dementia - a Review. *Dementia* (1994) 5(3-4):137-44. doi: Doi 10.1159/000106711. PubMed PMID: WOS:A1994NP58000004.

263. Rock PL, Roiser JP, Riedel WJ, Blackwell AD. Cognitive impairment in depression: a systematic review and meta-analysis. *Psychol Med* (2014) 44(10):2029-40. Epub 2013/10/31. doi: 10.1017/S0033291713002535. PubMed PMID: 24168753.

264. Schuch F, Vancampfort D, Firth J, Rosenbaum S, Ward P, Reichert T, et al. Physical activity and sedentary behavior in people with major depressive disorder: A systematic review and meta-analysis. *J Affect Disord* (2017) 210:139-50. Epub 2016/12/30. doi: 10.1016/j.jad.2016.10.050. PubMed PMID: 28033521.

265. Girard JM, Cohn JF, Mahoor MH, Mavadati SM, Hammal Z, Rosenwald DP. Nonverbal Social Withdrawal in Depression: Evidence from manual and automatic analysis. *Image Vis Comput* (2014) 32(10):641-7. Epub 2014/11/08. doi: 10.1016/j.imavis.2013.12.007. PubMed PMID: 25378765.

266. Smith DJ, Nicholl BI, Cullen B, Martin D, Ul-Haq Z, Evans J, et al. Prevalence and characteristics of probable major depression and bipolar disorder within UK biobank: cross-sectional study of 172,751 participants. *PLoS One* (2013) 8(11):e75362. doi: 10.1371/journal.pone.0075362. PubMed PMID: 24282498.

267. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. *Bmc Medical Research Methodology* (2017) 17. doi: ARTN 162; 10.1186/s12874-017-0442-1. PubMed PMID: WOS:000417509300002.

268. Lantz B. The large sample size fallacy. *Scandinavian Journal of Caring Sciences* (2013) 27(2):487-92. doi: 10.1111/j.1471-6712.2012.01052.x. PubMed PMID: WOS:000318815700034.

269. Ford C. Interpreting log transformations in a linear model: University of Virginia Library (2018) [cited 2020 12 Oct]. Available from: https://data.library.virginia.edu/interpreting-log-transformations-in-a-linear-model/.

270. Bakrania K, Edwardson CL, Khunti K, Bandelow S, Davies MJ, Yates T. Associations Between Sedentary Behaviors and Cognitive Function: Cross-Sectional and Prospective Findings From the UK Biobank. *Am J Epidemiol* (2018) 187(3):441-54. doi: 10.1093/aje/kwx273. PubMed PMID: WOS:000426812300008.

271. Leghsmith J, Wade DT, Hewer RL. Driving after a Stroke. *Journal of the Royal Society of Medicine* (1986) 79(4):200-3. doi: Doi 10.1177/014107688607900404. PubMed PMID: WOS:A1986A981200004.

272. Lings S, Jensen PB. Driving after stroke: a controlled laboratory investigation. *Int Disabil Stud* (1991) 13(3):74-82. Epub 1991/07/01. doi: 10.3109/03790799109166688. PubMed PMID: 1774218.

273. Fisk GD, Owsley C, Mennemeier M. Vision, attention, and self-reported driving behaviors in community-dwelling stroke survivors. *Arch Phys Med Rehabil* (2002) 83(4):469-77. Epub 2002/04/05. doi: 10.1053/apmr.2002.31179. PubMed PMID: 11932847.

274. Shiue I. Duration of daily TV/screen watching with cardiovascular, respiratory, mental and psychiatric health: Scottish Health Survey, 2012-2013. *Int J Cardiol* (2015) 186:241-6. Epub 2015/04/02. doi: 10.1016/j.ijcard.2015.03.259. PubMed PMID: 25828126.

275. Keadle SK, Moore SC, Sampson JN, Xiao Q, Albanes D, Matthews CE. Causes of Death Associated With Prolonged TV Viewing: NIH-AARP Diet and Health Study. *Am J Prev Med* (2015) 49(6):811-21. Epub 2015/07/29. doi: 10.1016/j.amepre.2015.05.023. PubMed PMID: 26215832.

276. Dempsey PC, Howard BJ, Lynch BM, Owen N, Dunstan DW. Associations of television viewing time with adults' well-being and vitality. *Preventive Medicine* (2014) 69:69-74. doi: 10.1016/j.ypmed.2014.09.007. PubMed PMID: WOS:000346221600013.

277. Fancourt D, Steptoe A. Television viewing and cognitive decline in older age: findings from the English Longitudinal Study of Ageing. *Sci Rep* (2019) 9(1):2851. Epub 2019/03/02. doi: 10.1038/s41598-019-39354-4. PubMed PMID: 30820029.

278. Wilson RS, Mendes De Leon CF, Barnes LL, Schneider JA, Bienias JL, Evans DA, et al. Participation in cognitively stimulating activities and risk of incident Alzheimer disease. *JAMA* (2002) 287(6):742-8. Epub 2002/02/20. doi: 10.1001/jama.287.6.742. PubMed PMID: 11851541.

279. Krell-Roesch J, Vemuri P, Pink A, Roberts RO, Stokin GB, Mielke MM, et al. Association Between Mentally Stimulating Activities in Late Life and the Outcome of Incident Mild Cognitive Impairment, With an Analysis of the APOE epsilon 4 Genotype. *Jama Neurology* (2017) 74(3):330-6. doi: 10.1001/jamaneurol.2016.3822. PubMed PMID: WOS:000397601900015.

280. Krell-Roesch J, Syrjanen JA, Vassilaki M, Machulda MM, Mielke MM, Knopman DS, et al. Quantity and quality of mental activities and the risk of incident mild cognitive impairment. *Neurology* (2019) 93(6):e548-e58. Epub 2019/07/12. doi: 10.1212/WNL.0000000000007897. PubMed PMID: 31292224.

281. Lampit A, Hallock H, Valenzuela M. Computerized Cognitive Training in Cognitively Healthy Older Adults: A Systematic Review and Meta-Analysis of Effect Modifiers. *Plos Medicine* (2014) 11(11). doi: ARTN e1001756; 10.1371/journal.pmed.1001756. PubMed PMID: WOS:000347001600005.

282. Lin JH, Lin RT, Tai CT, Hsieh CL, Hsiao SF, Liu CK. Prediction of poststroke dementia. *Neurology* (2003) 61(3):343-8. doi: Doi 10.1212/01.Wnl.0000078891.27052.10. PubMed PMID: WOS:000184712600013.

283. Potter GG, Helms MJ, Plassman BL. Associations of job demands and intelligence with cognitive performance among men in late life. *Neurology* (2008) 70(19):1803-8. doi: DOI 10.1212/01.wnl.0000295506.58497.7e. PubMed PMID: WOS:000256707200012.

284. Then FS, Luck T, Luppa M, Arelin K, Schroeter ML, Engel C, et al. Association between mental demands at work and cognitive functioning in the general population - results of the health study of the Leipzig research center for civilization diseases (LIFE). *J Occup Med Toxicol* (2014) 9. doi: Artn 23; 10.1186/1745-6673-9-23. PubMed PMID: WOS:000337528900001.

285. Smart EL, Gow AJ, Deary IJ. Occupational complexity and lifetime cognitive abilities. *Neurology* (2014) 83(24):2285-91. doi: 10.1212/Wnl.0000000000001075. PubMed PMID: WOS:000346070000019.

286. Prince SA, Adamo KB, Hamel ME, Hardt J, Gorber SC, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity* (2008) 5. doi: Artn 56; 10.1186/1479-5868-5-56. PubMed PMID: WOS:000261316900001.

287. Cheng ST, Chow PK, Song YQ, Yu ECS, Chan ACM, Lee TMC, et al. Mental and Physical Activities Delay Cognitive Decline in Older Persons with Dementia. *American Journal of Geriatric Psychiatry* (2014) 22(1):63-74. doi: 10.1016/j.jagp.2013.01.060. PubMed PMID: WOS:000336050800008.

288. Biddle SJH, Bennie JA, De Cocker K, Dunstan D, Gardiner PA, Healy GN, et al. Controversies in the Science of Sedentary Behaviour and Health: Insights, Perspectives and Future Directions from the 2018 Queensland Sedentary Behaviour Think Tank. *International Journal of Environmental Research and Public Health* (2019) 16(23). doi: ARTN 4762; 10.3390/ijerph16234762. PubMed PMID: WOS:000507275700069.

289. Hallgren M, Nguyen TTD, Owen N, Vancampfort D, Smith L, Dunstan DW, et al. Associations of interruptions to leisure-time sedentary behaviour with symptoms of depression and anxiety. *Translational Psychiatry* (2020) 10(1). doi: ARTN 128; 10.1038/s41398-020-0810-1. PubMed PMID: WOS:000533424100003.

290. Development and evaluation of strategies to reduce sedentary behaviour in patients after stroke and improve outcomes (RECREATE): Bradford Institute for Health Research [cited 2020 31 Oct]. Available from: https://www.bradfordresearch.nhs.uk/our-research-teams/academic-unit-for-ageing-and-stroke-research/our-research/stroke-research/recreate/.

291. Forster A, Birch KM, Carter G, Clarke DJ, English C, Farrin A, et al. The REduCing sedentaRy bEhaviour After sTrokE Study (RECREATE): Development and evaluation of an intervention to improve outcomes. *11th World Stroke Congress*; Montréal, Canada: Morressier (2018).

292. Doherty A, Jackson D, Hammerla N, Plotz T, Olivier P, Granat MH, et al. Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. *Plos One* (2017) 12(2). doi: ARTN e0169649; 10.1371/journal.pone.0169649. PubMed PMID: WOS:000396131700009.

293. Woodfield R, Sudlow CLM, Grp UBSO, Outcomes UBF-u. Accuracy of Patient Self-Report of Stroke: A Systematic Review from the UK Biobank Stroke Outcomes Group. *Plos One* (2015) 10(9). doi: ARTN e0137538; 10.1371/journal.pone.0137538. PubMed PMID: WOS:000360965800060.

294. King's College London. SSNAP Anual CCG and LHB Portfolio for April 2017 - March 2018 admissions and discharges. (2018).

295. Rundek T, Sacco RL. Prognosis after Stroke. In: Mohr JP, Wolf PA, Grotta JC, Moskowitz MA, Mayberg MR, von Kummer R, editors. *Stroke (Fifth Edition)*. Saint Louis: W.B. Saunders (2011). p. 219-41.

296. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* (2009) 338:b2393. Epub 2009/07/01. doi: 10.1136/bmj.b2393. PubMed PMID: 19564179.

297. Hawkley LC, Cacioppo JT. Loneliness matters: a theoretical and empirical review of consequences and mechanisms. *Ann Behav Med* (2010) 40(2):218-27. Epub 2010/07/24. doi: 10.1007/s12160-010-9210-8. PubMed PMID: 20652462.

298. Luo Y, Hawkley LC, Waite LJ, Cacioppo JT. Loneliness, health, and mortality in old age: a national longitudinal study. *Soc Sci Med* (2012) 74(6):907-14. Epub 2012/02/14. doi: 10.1016/j.socscimed.2011.11.028. PubMed PMID: 22326307.

299. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and Social Isolation as Risk Factors for Mortality: A Meta-Analytic Review. *Perspect Psychol Sci* (2015) 10(2):227-37. doi: 10.1177/1745691614568352. PubMed PMID: WOS:000351242300008.

300. Ong AD, Uchino BN, Wethington E. Loneliness and Health in Older Adults: A Mini-Review and Synthesis. *Gerontology* (2016) 62(4):443-9. doi: 10.1159/000441651. PubMed PMID: WOS:000378257900009.

301. Cacioppo JT, Hawkey LC. Perceived social isolation and cognition. *Trends Cogn Sci* (2009) 13(10):447-54. doi: 10.1016/j.tics.2009.06.005. PubMed PMID: WOS:000270656300005.

302. Tilvis RS, Kahonen-Vare MH, Jolkkonen J, Valvanne J, Pitkala KH, Strandberg TE. Predictors of cognitive decline and mortality of aged people over a 10-year period. *J Gerontol a-Biol* (2004) 59(3):268-74. PubMed PMID: WOS:000220368100010.

303. Wilson RS, Krueger KR, Arnold SE, Schneider JA, Kelly JF, Barnes LL, et al. Loneliness and risk of Alzheimer disease. *Arch Gen Psychiat* (2007) 64(2):234-40. doi: DOI 10.1001/archpsyc.64.2.234. PubMed PMID: WOS:000243990900011.

304. Seeman TE, Miller-Martinez DM, Merkin SS, Lachman ME, Tun PA, Karlamangla AS. Histories of Social Engagement and Adult Cognition: Midlife in the U.S. Study. *J Gerontol B-Psychol* (2011) 66:141-52. doi: 10.1093/geronb/gbq091. PubMed PMID: WOS:000293252000014.

305. Parisi JM, Stine-Morrow EAL, Noh SR, Morrow DG. Predispositional Engagement, Activity Engagement, and Cognition among Older Adults. *Aging, Neuropsychology, and Cognition* (2009) 16(4):485-504. doi: 10.1080/13825580902866653.

306. Lovden M, Ghisletta P, Lindenberger U. Social participation attenuates decline in perceptual speed in old and very old age. *Psychol Aging* (2005) 20(3):423-34. doi: 10.1037/0882-7974.20.3.423. PubMed PMID: WOS:000233022500006.

307. Ghisletta P, Bickel J-F, Lövdén M. Does Activity Engagement Protect Against Cognitive Decline in Old Age? Methodological and Analytical Considerations. *The Journals of Gerontology: Series B* (2006) 61(5):P253-P61. doi: 10.1093/geronb/61.5.P253.

308. Hosseini S, Chaurasia A, Oremus M. The Effect of Religion and Spirituality on Cognitive Function: A Systematic Review. *The Gerontologist* (2017) 59(2):e76-e85. doi: 10.1093/geront/gnx024.

309. Voas D, McAndrew S. Three Puzzles of Non-religion in Britain. *Journal of Contemporary Religion* (2012) 27(1):29-48. doi: 10.1080/13537903.2012.642725.

310. Carmelli D, Swan GE, LaRue A, Eslinger PJ. Correlates of change in cognitive function in survivors from the Western Collaborative Group Study. *Neuroepidemiology* (1997) 16(6):285-95. doi: Doi 10.1159/000109699. PubMed PMID: WOS:000071101600003.

311. Albert MS, Savage CR, Blazer D, Jones K, Berkman L, Seeman T, et al. Predictors of cognitive change in older persons: MacArthur studies of successful aging. *Psychol Aging* (1995) 10(4):578-89. doi: Doi 10.1037/0882-7974.10.4.578. PubMed PMID: WOS:A1995TL59300009.

312. Richards M, Hardy R, Wadsworth ME. Does active leisure protect cognition? Evidence from a national birth cohort. *Social science & medicine* (2003) 56(4):785-92.

313. Hatch SL, Feinstein L, Link BG, Wadsworth MEJ, RichardS M. The continuing benefits of education: Adult education and midlife cognitive ability in the British 1946 birth cohort. *J Gerontol B-Psychol* (2007) 62(6):S404-S14. doi: DOI 10.1093/geronb/62.6.S404. PubMed PMID: WOS:000253836400016.

314. Orrell M, Sahakian B. Education and Dementia - Research Evidence Supports the Concept Use It or Lose It. *Bmj-Brit Med J* (1995) 310(6985):951-2. doi: DOI 10.1136/bmj.310.6985.951. PubMed PMID: WOS:A1995QT90400001.

315. Ellwardt L, Aartsen M, Deeg D, Steverink N. Does loneliness mediate the relation between social support and cognitive functioning in later life? *Social Science & Medicine* (2013) 98:116-24. doi: https://doi.org/10.1016/j.socscimed.2013.09.002.

316. Byrne C, Saville CWN, Coetzer R, Ramsey R. Stroke survivors experience elevated levels of loneliness: a multi-year analysis of the National Survey for Wales. *medRxiv* (2020):2020.07.02.20145417. doi: 10.1101/2020.07.02.20145417.

317. Department for Digital Culture, Media & Sport, Prime Minister's Office. *£20 million investment to help tackle loneliness*. (2018).

318. Masi CM, Chen HY, Hawkley LC, Cacioppo JT. A Meta-Analysis of Interventions to Reduce Loneliness. *Pers Soc Psychol Rev* (2011) 15(3):219-66. doi: 10.1177/1088868310377394. PubMed PMID: WOS:000292207700001.

319. Dickens AP, Richards SH, Greaves CJ, Campbell JL. Interventions targeting social isolation in older people: a systematic review. *Bmc Public Health* (2011) 11. doi: Artn 647; 10.1186/1471-2458-11-647. PubMed PMID: WOS:000294774000001.

320. Theeke LA, Mallow JA. The Development of LISTEN: A Novel Intervention for Loneliness. *Open J Nurs* (2015) 5(2):136-43. Epub 2015/08/01. doi: 10.4236/ojn.2015.52016. PubMed PMID: 26229740.

321. Theeke LA, Mallow JA, Barnes ER, Theeke E. The Feasibility and Acceptability of LISTEN for Loneliness. *Open J Nurs* (2015) 5(5):416-25. Epub 2015/09/25. doi: 10.4236/ojn.2015.55045. PubMed PMID: 26401420.

322. Theeke L, Lucke-Wold AN, Mallow J, Horstman P. Life after stroke in Appalachia. *Int J Nurs Sci* (2017) 4(2):105-11. Epub 2017/03/10. doi: 10.1016/j.ijnss.2017.02.005. PubMed PMID: 31406728.

323. NHS Improvement - Stroke. Psychological care after stroke - Improving stroke services for people with cognitive and mood disorders.(2011. Available from: https://www.nice.org.uk/media/default/sharedlearning/531_strokepsychologicalsupportfinal.pdf.

324. Lincoln KD. Social Support, Negative Social Interactions, and Psychological Well-Being. *Soc Serv Rev* (2000) 74(2):231-52. Epub 2000/06/01. doi: 10.1086/514478. PubMed PMID: 26594064.

325. Cohen S. Social relationships and health. *Am Psychol* (2004) 59(8):676-84. Epub 2004/11/24. doi: 10.1037/0003-066X.59.8.676. PubMed PMID: 15554821.

326. Thompson SC, Sobolew-Shubin A. Overprotective Relationships - a Nonsupportive Side of Social Networks. *Basic Appl Soc Psych* (1993) 14(3):363-83. doi: DOI 10.1207/s15324834basp1403_8. PubMed PMID: WOS:A1993LW48100008.

327. Croteau C, Le Dorze G. Overprotection, "speaking for", and conversational participation: A study of couples with aphasia. *Aphasiology* (2006) 20(2-4):327-36. doi: 10.1080/02687030500475051. PubMed PMID: WOS:000237890000019.

328. Bailey R. Examining daily physical activity in community-dwelling adults with stroke using social cognitive theory: an exploratory, qualitative study. *Disabil Rehabil* (2020) 42(18):2631-9. doi: 10.1080/09638288.2019.1568591. PubMed PMID: WOS:000464631800001.

329. Haun J, Rittman M, Sberna M. The continuum of connectedness and social isolation during post stroke recovery. *J Aging Stud* (2008) 22(1):54-64. Epub 2008/01/01. doi: 10.1016/j.jaging.2007.03.001. PubMed PMID: 26047096.

330. Northcott S, Hilari K. Why do people lose their friends after a stroke? *International Journal of Language & Communication Disorders* (2011) 46(5):524-34. doi: 10.1111/j.1460-6984.2011.00079.x. PubMed PMID: WOS:000295872600003.

331. Pels F, Kleinert J. Loneliness and physical activity: A systematic review. *Int Rev Sport Exer P* (2016) 9(1):231-60. doi: 10.1080/1750984x.2016.1177849. PubMed PMID: WOS:000382261400011.

332. Drozdowska BA, Elliott E, Taylor-Rowan M, Shaw RC, Cuthbertson G, Langhorne P, et al. Cardiovascular risk factors indirectly affect acute post-stroke cognition through stroke severity and prior cognitive impairment: a moderated mediation analysis. *Alzheimer's Research & Therapy* (2020) 12(1):85. doi: 10.1186/s13195-020-00653-y.

333. Yang ZR, Wang HYY, Edwards D, Ding CY, Yan L, Brayne C, et al. Association of blood lipids, atherosclerosis and statin use with dementia and cognitive impairment after stroke: A systematic review and meta-analysis. *Ageing Res Rev* (2020) 57. doi: ARTN 100962; 10.1016/j.arr.2019.100962. PubMed PMID: WOS:000509630900012.

334. Gorelick PB. Risk factors for vascular dementia and Alzheimer disease. *Stroke* (2004) 35(11 Suppl 1):2620-2. Epub 2004/09/18. doi: 10.1161/01.STR.0000143318.70292.47. PubMed PMID: 15375299.

335. Sahathevan R, Brodtmann A, Donnan GA. Dementia, stroke, and vascular risk factors; a review. *Int J Stroke* (2012) 7(1):61-73. Epub 2011/12/23. doi: 10.1111/j.1747-4949.2011.00731.x. PubMed PMID: 22188853.

336. Newman AB, Fitzpatrick AL, Lopez O, Jackson S, Lyketsos C, Jagust W, et al. Dementia and Alzheimer's disease incidence in relationship to cardiovascular disease in the Cardiovascular Health Study cohort. *J Am Geriatr Soc* (2005) 53(7):1101-7.

337. Leritz EC, McGlinchey RE, Kellison I, Rudolph JL, Milberg WP. Cardiovascular Disease Risk Factors and Cognition in the Elderly. *Curr Cardiovasc Risk Rep* (2011) 5(5):407-12. doi: 10.1007/s12170-011-0189-x. PubMed PMID: 22199992.

338. Vakhnina NV, Nikitina LY, Parfenov VA, Yakhno NN. Post-stroke cognitive impairments. *Neurosci Behav Physiol* (2009) 39(8):719-24. Epub 2009/09/26. doi: 10.1007/s11055-009-9198-3. PubMed PMID: 19779824.

339. Henon H, Durieu I, Guerouaou D, Lebert F, Pasquier F, Leys D. Poststroke dementia: incidence and relationship to prestroke cognitive decline. *Neurology* (2001) 57(7):1216-22. Epub 2001/10/10. doi: 10.1212/wnl.57.7.1216. PubMed PMID: 11591838.

340. Bunch TJ, Weiss JP, Crandall BG, May HT, Bair TL, Osborn JS, et al. Atrial fibrillation is independently associated with senile, vascular, and Alzheimer's dementia. *Heart rhythm* (2010) 7(4):433-7.

341. Jørgensen HS, Nakayama H, Reith J, Raaschou HO, Olsen TS. Acute stroke with atrial fibrillation: the Copenhagen Stroke Study. *Stroke* (1996) 27(10):1765-9.

342. Kirino T. Ischemic tolerance. *J Cereb Blood Flow Metab* (2002) 22(11):1283-96. Epub 2002/11/20. doi: 10.1097/01.WCB.0000040942.89393.88. PubMed PMID: 12439285.

343. Weih M, Kallenberg K, Bergk A, Dirnagl U, Harms L, Wernecke K, et al. Attenuated stroke severity after prodromal TIA: a role for ischemic tolerance in the brain? *stroke* (1999) 30(9):1851-4.

344. Moncayo J, de Freitas GR, Bogousslavsky J, Altieri M, van Melle G. Do transient ischemic attacks have a neuroprotective effect? *Neurology* (2000) 54(11):2089-94. doi: 10.1212/wnl.54.11.2089.

345. Wegener S, Gottschalk B, Jovanovic V, Knab R, Fiebach JB, Schellinger PD, et al. Transient ischemic attacks before ischemic stroke: preconditioning the human brain? A multicenter magnetic resonance imaging study. *Stroke* (2004) 35(3):616-21.

346. Nishijima Y, Akamatsu Y, Weinstein PR, Liu J. Collaterals: Implications in cerebral ischemic diseases and therapeutic interventions. *Brain Res* (2015) 1623:18-29. Epub 2015/03/17. doi: 10.1016/j.brainres.2015.03.006. PubMed PMID: 25770816.

347. Liebeskind DS, Cotsonis GA, Saver JL, Lynn MJ, Cloft HJ, Chimowitz MI, et al. Collateral circulation in symptomatic intracranial atherosclerosis. *J Cereb Blood Flow Metab* (2011) 31(5):1293-301. Epub 2010/12/16. doi: 10.1038/jcbfm.2010.224. PubMed PMID: 21157476.

348. Romero JR, Pikula A, Nguyen TN, Nien YL, Norbash A, Babikian VL. Cerebral collateral circulation in carotid artery disease. *Curr Cardiol Rev* (2009) 5(4):279-88. doi: 10.2174/157340309789317887. PubMed PMID: 21037845.

349. Shuaib A, Butcher K, Mohammad AA, Saqqur M, Liebeskind DS. Collateral blood vessels in acute ischaemic stroke: a potential therapeutic target. *Lancet Neurol* (2011) 10(10):909-21. Epub 2011/09/24. doi: 10.1016/S1474-4422(11)70195-8. PubMed PMID: 21939900.

350. Wufuer A, Wubuli A, Mijiti P, Zhou J, Tuerxun S, Cai J, et al. Impact of collateral circulation status on favorable outcomes in thrombolysis treatment: A systematic review and meta-analysis. *Exp Ther Med* (2018) 15(1):707-18. Epub 2018/02/06. doi: 10.3892/etm.2017.5486. PubMed PMID: 29399075.

351. Liebeskind DS. Collateral circulation. *Stroke* (2003) 34(9):2279-84.

352. Hedera P, Bujdakova J, Traubner P, Pancak J. Stroke risk factors and development of collateral flow in carotid occlusive disease. *Acta Neurol Scand* (1998) 98(3):182-6. Epub 1998/10/24. doi: 10.1111/j.1600-0404.1998.tb07291.x. PubMed PMID: 9786615.

353. Ergul A, Abdelsaid M, Fouda AY, Fagan SC. Cerebral Neovascularization in Diabetes: Implications for Stroke Recovery and beyond. *Journal of Cerebral Blood Flow & Metabolism* (2014) 34(4):553-63. doi: 10.1038/jcbfm.2014.18.

354. Rizk H, Allam M, Hegazy A, Khalil H, Helmy H, Hashem HS, et al. Predictors of poor cerebral collaterals and cerebrovascular reserve in patients with chronic total carotid occlusion. *International Journal of Neuroscience* (2019) 129(5):455-60.

355. Pires PW, Dams Ramos CM, Matin N, Dorrance AM. The effects of hypertension on the cerebral circulation. *Am J Physiol Heart Circ Physiol* (2013) 304(12):H1598-614. Epub 2013/04/16. doi: 10.1152/ajpheart.00490.2012. PubMed PMID: 23585139.

356. Ferdinandy P, Hausenloy DJ, Heusch G, Baxter GF, Schulz R. Interaction of risk factors, comorbidities, and comedications with ischemia/reperfusion injury and cardioprotection by preconditioning, postconditioning, and remote conditioning. *Pharmacological reviews* (2014) 66(4):1142-74.

357. Zheng Y, Wang M. Clinical study about effects of the acute fasting hyperglycemia and type 2 diabetes mellitus on ischemic preconditioning in patients with the first-ever acute anterior circulation infarction. *Chinese Journal of Neurology* (2008) 41(12):820-3.

358. Fairchild AJ, McDaniel HL. Best (but oft-forgotten) practices: mediation analysis. *Am J Clin Nutr* (2017) 105(6):1259-71. Epub 2017/04/28. doi: 10.3945/ajcn.117.152546. PubMed PMID: 28446497.

359. Hayes AF. *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*: Guilford Publications (2017).

360. Kasner SE, Chalela JA, Luciano JM, Cucchiara BL, Raps EC, McGarvey ML, et al. Reliability and validity of estimating the NIH stroke scale score from medical records. *Stroke* (1999) 30(8):1534-7. Epub 1999/08/06. doi: 10.1161/01.str.30.8.1534. PubMed PMID: 10436096.

361. Fussner J, Velasco C. Stroke coordinator boot camp: Assessing stroke - scores & scales. [PowerPoint Presentation]: American Stroke Association [cited 2019 October 15]. Available from: https://www.heart.org/-/media/files/affiliates/gra/gra-qsi/2019-scbc-presentations/5--assessing-stroke--scores--scales-v2.pdf?la=en&hash=ED3F9267A585CFDE7514E8D9A8180B07E0BECA6F.

362. World Health Organization (WHO). International Classification of Diseases for Mortality and Morbidity Statistics (11th ed.) (2019). Available from: https://icd.who.int/.

363. Hodkinson HM. Evaluation of a mental test score for assessment of mental impairment in the elderly. *Age Ageing* (1972) 1(4):233-8. Epub 1972/11/01. doi: 10.1093/ageing/1.4.233. PubMed PMID: 4669880.

364. Abbreviated Mental Test Score (AMTS). *Occas Pap R Coll Gen Pract* (1993) (59):28. PubMed PMID: 19790838.

365. Lees RA, Hendry Ba K, Broomfield N, Stott D, Larner AJ, Quinn TJ. Cognitive assessment in stroke: feasibility and test properties using differing approaches to scoring of incomplete items. *Int J Geriatr Psychiatry* (2017) 32(10):1072-8. Epub 2016/08/17. doi: 10.1002/gps.4568. PubMed PMID: 27526678.

366. Muthén LK, Muthén BO. *Mplus User's Guide. Eighth Edition*. Los Angeles, CA: Muthén & Muthén ((1998-2017)).

367. Asparouhov T, Muthén B. Weighted least squares estimation with missing data. *Mplus Technical Appendix* (2010) 2010:1-10.

368. Aiken LS, West SG, Reno RR. *Multiple regression: Testing and interpreting interactions*: Sage (1991).

369. Wright S. Correlation and causation Part I. Method of path coefficients. *J Agric Res* (1920) 20:0557-85. PubMed PMID: WOS:000188347400034.

370. Hayes AF. Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation. *Commun Monogr* (2018) 85(1):4-40. doi: 10.1080/03637751.2017.1352100. PubMed PMID: WOS:000428291600002.

371. Shrout PE, Bolger N. Mediation in experimental and nonexperimental studies: new procedures and recommendations. *Psychol Methods* (2002) 7(4):422-45. Epub 2003/01/18. PubMed PMID: 12530702.

372. Mallinckrodt B, Abraham WT, Wei MF, Russell DW. Advances in testing the statistical significance of mediation effects. *J Couns Psychol* (2006) 53(3):372-8. doi: 10.1037/0022-0167.53.3.372. PubMed PMID: WOS:000239045900009.

373. Fritz MS, Mackinnon DP. Required sample size to detect the mediated effect. *Psychol Sci* (2007) 18(3):233-9. Epub 2007/04/21. doi: 10.1111/j.1467-9280.2007.01882.x. PubMed PMID: 17444920.

374. Preacher KJ, Rucker DD, Hayes AF. Addressing Moderated Mediation Hypotheses: Theory, Methods, and Prescriptions. *Multivariate Behav Res* (2007) 42(1):185-227. Epub 2007/01/01. doi: 10.1080/00273170701341316. PubMed PMID: 26821081.

375. Ditlevsen S, Christensen U, Lynch J, Damsgaard MT, Keiding N. The mediation proportion - A structural equation approach for estimating the proportion of exposure effect on outcome explained by an intermediate variable. *Epidemiology* (2005) 16(1):114-20. doi: 10.1097/01.ede.0000147107.76079.07. PubMed PMID: WOS:000226079600017.

376. Fairchild AJ, McQuillin SD. Evaluating mediation and moderation effects in school psychology: a presentation of methods and review of current practice. *J Sch Psychol* (2010) 48(1):53-84. Epub 2009/12/17. doi: 10.1016/j.jsp.2009.09.001. PubMed PMID: 20006988.

377. Hooper D, Coughlan J, Mullen M. Structural equation modelling: guidelines for determining model fit. *The Electronic Journal of Business Research Methods* (2008) 6(1):53-60.

378. Manabe Y, Kono S, Tanaka T, Narai H, Omori N. High blood pressure in acute ischemic stroke and clinical outcome. *Neurol Int* (2009) 1(1):e1. Epub 2009/01/01. doi: 10.4081/ni.2009.e1. PubMed PMID: 21577346.

379. Willmot M, Leonardi-Bee J, Bath PM. High blood pressure in acute stroke and subsequent outcome: a systematic review. *Hypertension* (2004) 43(1):18-24. Epub 2003/12/10. doi: 10.1161/01.HYP.0000105052.65787.35. PubMed PMID: 14662649.

380. Hong L, Cheng X, Lin L, Bivard A, Ling Y, Butcher K, et al. The blood pressure paradox in acute ischemic stroke. *Ann Neurol* (2019) 85(3):331-9. Epub 2019/02/06. doi: 10.1002/ana.25428. PubMed PMID: 30720216.

381. Havranek EP, Mujahid MS, Barr DA, Blair IV, Cohen MS, Cruz-Flores S, et al. Social Determinants of Risk and Outcomes for Cardiovascular Disease: A Scientific Statement From the American Heart Association. *Circulation* (2015) 132(9):873-98. Epub 2015/08/05. doi: 10.1161/CIR.0000000000000228. PubMed PMID: 26240271.

382. Johnston SC. Ischemic preconditioning from transient ischemic attacks? Data from the Northern California TIA Study. *Stroke* (2004) 35(11 Suppl 1):2680-2. Epub 2004/09/25. doi: 10.1161/01.STR.0000143322.20491.0f. PubMed PMID: 15388902.

383. Della Morte D, Abete P, Gallucci F, Scaglione A, D'Ambrosio D, Gargiulo G, et al. Transient ischemic attack before nonlacunar ischemic stroke in the elderly. *J Stroke Cerebrovasc Dis* (2008) 17(5):257-62. Epub 2008/08/30. doi: 10.1016/j.jstrokecerebrovasdis.2008.03.004. PubMed PMID: 18755403.

384. Correia PN, Meyer IA, Eskandari A, Amiguet M, Hirt L, Michel P. Preconditioning by Preceding Ischemic Cerebrovascular Events. *J Am Heart Assoc* (2021) 10(16):e020129. Epub 2021/08/14. doi: 10.1161/JAHA.120.020129. PubMed PMID: 34387096.

385. Hao Y, Xin M, Feng L, Wang X, Wang X, Ma D, et al. Review Cerebral Ischemic Tolerance and Preconditioning: Methods, Mechanisms, Clinical Applications, and Challenges. *Front Neurol* (2020) 11:812. Epub 2020/10/20. doi: 10.3389/fneur.2020.00812. PubMed PMID: 33071923.

386. Rothwell PM, Algra A, Chen Z, Diener HC, Norrving B, Mehta Z. Effects of aspirin on risk and severity of early recurrent stroke after transient ischaemic attack and ischaemic stroke: time-course analysis of randomised trials. *Lancet* (2016) 388(10042):365-75. Epub 2016/05/23. doi: 10.1016/S0140-6736(16)30468-8. PubMed PMID: 27209146.

387. Sargento-Freitas J, Pagola J, Rubiera M, Flores A, Silva F, Rodriguez-Luna D, et al. Preferential Effect of Premorbid Statins on Atherothrombotic Strokes through Collateral Circulation Enhancement. *Eur Neurol* (2012) 68(3):171-6. doi: 10.1159/000337862. PubMed PMID: WOS:000308607800009.

388. Ovbiagele B, Saver JL, Starkman S, Kim D, Ali LK, Jahan R, et al. Statin enhancement of collateralization in acute stroke. *Neurology* (2007) 68(24):2129-31. Epub 2007/06/15. doi: 10.1212/01.wnl.0000264931.34941.f0. PubMed PMID: 17562834.

389. Lang L, Clifford A, Wei L, Zhang D, Leung D, Augustine G, et al. Prevalence and determinants of undetected dementia in the community: a systematic literature review and a meta-analysis. *BMJ Open* (2017) 7(2):e011146. Epub 2017/02/06. doi: 10.1136/bmjopen-2016-011146. PubMed PMID: 28159845.

390. Oh H, Park J, Seo W. A 2-year prospective follow-up study of temporal changes associated with post-stroke cognitive impairment. *Int J Nurs Pract* (2018) 24(2):e12618. Epub 2018/01/02. doi: 10.1111/ijn.12618. PubMed PMID: 29291599.

391. del Ser T, Barba R, Morin MM, Domingo J, Cemillan C, Pondal M, et al. Evolution of cognitive impairment after stroke and risk factors for delayed progression. *Stroke* (2005) 36(12):2670-5. Epub 2005/10/29. doi: 10.1161/01.STR.0000189626.71033.35. PubMed PMID: 16254227.

392. Nys GM, van Zandvoort MJ, de Kort PL, Jansen BP, de Haan EH, Kappelle LJ. Cognitive disorders in acute stroke: prevalence and clinical determinants. *Cerebrovasc Dis* (2007) 23(5-6):408-16. Epub 2007/04/05. doi: 10.1159/000101464. PubMed PMID: 17406110.

393. Goodglass H, Kaplan E. *The assessment of aphasia and related disorders*. Philadelphia,: Lea & Febiger (1972).

394. Benton A, Hamsher K. Multilingual aphasia examination manual. *Iowa City: University of Iowa* (1978).

395. Jolles J, Houx PJ, van Boxtel MPJ, Ponds RWHM. The Maastricht Aging Study: Determinants of cognitive aging. Maastricht, The Netherlands: Neuropsych Publishers Maastricht (1995).

396. Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* (1989) 39(9):1159-65. Epub 1989/09/01. doi: 10.1212/wnl.39.9.1159. PubMed PMID: 2771064.

397. *Army Individual Test Battery. Manual of directions and scoring*. Washington, DC: War Department, Adjutant General's Office (1944).

398. Reitan RM, Wolfson D. *The Halstead-Reitan Neuropsychological Test Battery: Therapy and clinical interpretation*. Tucson, AZ: Neuropsychological Press (1985).

399. Tombaugh TN. Trail Making Test A and B: normative data stratified by age and education. *Arch Clin Neuropsychol* (2004) 19(2):203-14. Epub 2004/03/11. doi: 10.1016/S0887-6177(03)00039-8. PubMed PMID: 15010086.

400. Kaplan EF, Goodglass H, Weintraub S. *The Boston Naming Test*. Boston, MA: Veterans Administration Medical Center (1978).

401. Rey A. L'examen psychologique dans les cas d'encephalopathie traumatique. *Archives de Psychologie* (1941) 28:286-340.

402. Osterrieth PA. Le test de copie d'une figure complexe; contribution à l'étude de la perception et de la mémoire. *Archives de Psychologie* (1944) 30:206-356.

403. Shin MS, Park SY, Park SR, Seol SH, Kwon JS. Clinical and empirical applications of the Rey-Osterrieth Complex Figure Test. *Nat Protoc* (2006) 1(2):892-9. doi: 10.1038/nprot.2006.115. PubMed PMID: WOS:000251002500052.

404. Mahoney FI, Barthel DW. Functional Evaluation: The Barthel Index. *Maryland State Medical Journal* (1965) 14:61-5. Epub 1965/02/01. PubMed PMID: 14258950.

405. Cesari M, Demougeot L, Boccalon H, Guyonnet S, Abellan Van Kan G, Vellas B, et al. A self-reported screening tool for detecting community-dwelling older persons with frailty syndrome in the absence of mobility disability: the FiND questionnaire. *PLoS One* (2014) 9(7):e101745. Epub 2014/07/08. doi: 10.1371/journal.pone.0101745. PubMed PMID: 24999805.

406. Mendoza TR, Wang XS, Cleeland CS, Morrissey M, Johnson BA, Wendt JK, et al. The rapid assessment of fatigue severity in cancer patients: use of the Brief Fatigue Inventory. *Cancer* (1999) 85(5):1186-96. Epub 1999/03/26. doi: 10.1002/(sici)1097-0142(19990301)85:5<1186::aid-cncr24>3.0.co;2-n. PubMed PMID: 10091805.

407. Marshall AL, Smith BJ, Bauman AE, Kaur S. Reliability and validity of a brief physical activity assessment for use by family doctors. *British Journal of Sports Medicine* (2005) 39(5):294-7. doi: 10.1136/bjsm.2004.013771. PubMed PMID: WOS:000228635600011.

408. Jenkinson C, Fitzpatrick R, Crocker H, Peters M. The Stroke Impact Scale: validation in a UK setting and development of a SIS short form and SIS index. *Stroke* (2013) 44(9):2532-5. Epub 2013/07/23. doi: 10.1161/STROKEAHA.113.001847. PubMed PMID: 23868278.

409. Turner-Stokes L, Kalmus M, Hirani D, Clegg F. The Depression Intensity Scale Circles (DISCs): a first evaluation of a simple assessment tool for depression in the context of brain injury. *J Neurol Neurosurg Psychiatry* (2005) 76(9):1273-8. Epub 2005/08/19. doi: 10.1136/jnnp.2004.050096. PubMed PMID: 16107367.

410. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. *Med Care* (2003) 41(11):1284-92. Epub 2003/10/30. doi: 10.1097/01.MLR.0000093487.78664.3C. PubMed PMID: 14583691.

411. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Lowe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med* (2007) 146(5):317-25. Epub 2007/03/07. doi: 10.7326/0003-4819-146-5-200703060-00004. PubMed PMID: 17339617.

412. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* (2001) 16(9):606-13. Epub 2001/09/15. doi: 10.1046/j.1525-1497.2001.016009606.x. PubMed PMID: 11556941.

413. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* (2006) 166(10):1092-7. Epub 2006/05/24. doi: 10.1001/archinte.166.10.1092. PubMed PMID: 16717171.

414. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med* (2002) 64(2):258-66. Epub 2002/03/27. doi: 10.1097/00006842-200203000-00008. PubMed PMID: 11914441.

415. Kroenke K, Spitzer RL, Williams JB, Lowe B. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. *Gen Hosp Psychiatry* (2010) 32(4):345-59. Epub 2010/07/17. doi: 10.1016/j.genhosppsych.2010.03.006. PubMed PMID: 20633738.

416. Eaton WW, Smith C, Ybarra M, Muntaner C, Tien A. Center for Epidemiologic Studies Depression Scale: Review and Revision (CESD and CESD-R). *The use of psychological testing for treatment planning and outcomes assessment: Instruments for adults, Volume 3, 3rd ed*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers (2004). p. 363-77.

417. EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy* (1990) 16(3):199-208. Epub 1990/11/05. doi: 10.1016/0168-8510(90)90421-9. PubMed PMID: 10109801.

418. Gjesfjeld CD, Greeno CG, Kim KH. A confirmatory factor analysis of an abbreviated social support instrument: The MOS-SSS. *Res Social Work Prac* (2008) 18(3):231-7. doi: 10.1177/1049731507309830. PubMed PMID: WOS:000254983900006.

419. Patchick E, Vail A, Wood A, Bowen A. PRECiS (Patient Reported Evaluation of Cognitive State): psychometric evaluation of a new patient reported outcome measure of the impact of stroke. *Clin Rehabil* (2016) 30(12):1229-41. Epub 2016/01/02. doi: 10.1177/0269215515624480. PubMed PMID: 26721872.

420. Bamford J, Sandercock P, Dennis M, Burn J, Warlow C. Classification and natural history of clinically identifiable subtypes of cerebral infarction. *Lancet* (1991) 337(8756):1521-6. Epub 1991/06/22. doi: 10.1016/0140-6736(91)93206-o. PubMed PMID: 1675378.

421. Brandt J, Spencer M, Folstein M. The telephone interview for cognitive status. *Neuropsychiatry Neuropsychol Behav Neurol* (1988) 1(2):111-7.

422. Brandt J, Welsh KA, Breitner JC, Folstein MF, Helms M, Christian JC. Hereditary influences on cognitive functioning in older men. A study of 4000 twin pairs. *Arch Neurol* (1993) 50(6):599-603. Epub 1993/06/01. doi: 10.1001/archneur.1993.00540060039014. PubMed PMID: 8503796.

423. Jorm AF. A short form of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): development and cross-validation. *Psychol Med* (1994) 24(1):145-53. Epub 1994/02/01. doi: 10.1017/s003329170002691x. PubMed PMID: 8208879.

424. Galvin JE, Roe CM, Powlishta KK, Coats MA, Muich SJ, Grant E, et al. The AD8: a brief informant interview to detect dementia. *Neurology* (2005) 65(4):559-64. Epub 2005/08/24. doi: 10.1212/01.wnl.0000172958.95282.2a. PubMed PMID: 16116116.

425. Bennett HE, Thomas SA, Austen R, Morris AM, Lincoln NB. Validation of screening measures for assessing mood in stroke patients. *Br J Clin Psychol* (2006) 45(Pt 3):367-76. Epub 2006/12/07. doi: 10.1348/014466505x58277. PubMed PMID: 17147102.

426. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clinical Gerontologist: The Journal of Aging and Mental Health* (1986).

427. Kaufer DI, Cummings JL, Ketchel P, Smith V, MacMillan A, Shelley T, et al. Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory. *J Neuropsychiatry Clin Neurosci* (2000) 12(2):233-9. Epub 2000/09/23. doi: 10.1176/jnp.12.2.233. PubMed PMID: 11001602.

428. Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist* (1980) 20(6):649-55. Epub 1980/12/01. doi: 10.1093/geront/20.6.649. PubMed PMID: 7203086.

429. Sentinel Stroke National Audit Programme: King's College London [cited 2021 August 22]. Available from: https://www.strokeaudit.org/About-SSNAP.aspx.

430. Muthen B, Muthen LK. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol Clin Exp Res* (2000) 24(6):882-91. Epub 2000/07/11. PubMed PMID: 10888079.

431. Desmond DW, Moroney JT, Sano M, Stern Y. Recovery of cognitive function after stroke. *Stroke* (1996) 27(10):1798-803. Epub 1996/10/01. doi: 10.1161/01.str.27.10.1798. PubMed PMID: 8841333.

432. Hochstenbach JB, den Otter R, Mulder TW. Cognitive recovery after stroke: a 2-year follow-up. *Arch Phys Med Rehabil* (2003) 84(10):1499-504. Epub 2003/10/31. doi: 10.1016/s0003-9993(03)00370-8. PubMed PMID: 14586918.

433. de Haan EH, Nys GM, Van Zandvoort MJ. Cognitive function following stroke and vascular cognitive impairment. *Curr Opin Neurol* (2006) 19(6):559-64. Epub 2006/11/15. doi: 10.1097/01.wco.0000247612.21235.d9. PubMed PMID: 17102694.

434. Srikanth VK, Quinn SJ, Donnan GA, Saling MM, Thrift AG. Long-term cognitive transitions, rates of cognitive change, and predictors of incident dementia in a population-based first-ever stroke cohort. *Stroke* (2006) 37(10):2479-83. Epub 2006/09/02. doi: 10.1161/01.STR.0000239666.46828.d7. PubMed PMID: 16946165.

435. Patel M, Coshall C, Rudd AG, Wolfe CD. Natural history of cognitive impairment after stroke and factors associated with its recovery. *Clin Rehabil* (2003) 17(2):158-66. Epub 2003/03/11. doi: 10.1191/0269215503cr596oa. PubMed PMID: 12625656.

436. Liman TG, Heuschmann PU, Endres M, Floel A, Schwab S, Kolominsky-Rabas PL. Changes in cognitive function over 3 years after first-ever stroke and predictors of cognitive impairment and long-term cognitive stability: the Erlangen Stroke Project. *Dement Geriatr Cogn Disord* (2011) 31(4):291-9. Epub 2011/04/20. doi: 10.1159/000327358. PubMed PMID: 21502760.

437. Muthen B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* (1999) 55(2):463-9. doi: DOI 10.1111/j.0006-341X.1999.00463.x. PubMed PMID: WOS:000081320800018.

438. Muthen B, Asparouhov T. Growth mixture modeling: Analysis with non-Gaussian random effects. *Ch Crc Handb Mod Sta* (2009):143-65. PubMed PMID: WOS:000267748700006.

439. Nagin DS, Land KC. Age, Criminal Careers, and Population Heterogeneity - Specification and Estimation of a Nonparametric, Mixed Poisson Model. *Criminology* (1993) 31(3):327-62. doi: DOI 10.1111/j.1745-9125.1993.tb01133.x. PubMed PMID: WOS:A1993LQ01600002.

440. Jung T, Wickrama KA. An introduction to latent class growth analysis and growth mixture modeling. *Social and personality psychology compass* (2008) 2(1):302-17.

441. van de Schoot R, Sijbrandij M, Winter SD, Depaoli S, Vermunt JK. The GRoLTS-Checklist: Guidelines for Reporting on Latent Trajectory Studies. *Struct Equ Modeling* (2017) 24(3):451-67. doi: 10.1080/10705511.2016.1247646. PubMed PMID: WOS:000398193800010.

442. van Mierlo M, van Heugten C, Post MWM, Hoekstra T, Visser-Meily A. Trajectories of health-related quality of life after stroke: results from a one-year prospective cohort study. *Disabil Rehabil* (2018) 40(9):997-1006. doi: 10.1080/09638288.2017.1292320. PubMed PMID: WOS:000425689200002.

443. Coulombe P, Selig JP, Delaney HD. Ignoring individual differences in times of assessment in growth curve modeling. *Int J Behav Dev* (2016) 40(1):76-86. doi: 10.1177/0165025415577684. PubMed PMID: WOS:000366603500009.

444. Berlin KS, Parra GR, Williams NA. An Introduction to Latent Variable Mixture Modeling (Part 2): Longitudinal Latent Class Growth Analysis and Growth Mixture Models. *J Pediatr Psychol* (2014) 39(2):188-203. doi: 10.1093/jpepsy/jst085. PubMed PMID: WOS:000332344000007.

445. Curran PJ, Obeidat K, Losardo D. Twelve Frequently Asked Questions About Growth Curve Modeling. *J Cogn Dev* (2010) 11(2):121-36. doi: Pii 921688997; 10.1080/15248371003699969. PubMed PMID: WOS:000277567500001.

446. Shi D, DiStefano C, Zheng X, Liu R, Jiang Z. Fitting Latent Growth Models with Small Sample Sizes and Non-normal Missing Data. *Int J Behav Dev* (2021) 45(2):179-92. Epub 2021/03/06. doi: 10.1177/0165025420979365. PubMed PMID: 33664535.

447. McNeish D, Wolf MG. Thinking twice about sum scores. *Behav Res Methods* (2020) 52(6):2287-305. doi: 10.3758/s13428-020-01398-0. PubMed PMID: WOS:000528112600001.

448. Freitas S, Prieto G, Simoes MR, Santana I. Psychometric properties of the Montreal Cognitive Assessment (MoCA): an analysis using the Rasch model. *Clin Neuropsychol* (2014) 28(1):65-83. Epub 2014/01/11. doi: 10.1080/13854046.2013.870231. PubMed PMID: 24404822.

449. Dimitrov DM. Testing for Factorial Invariance in the Context of Construct Validation. *Meas Eval Couns Dev* (2010) 43(2):121-49. doi: 10.1177/0748175610373459. PubMed PMID: WOS:000278947000004.

450. Millsap RE, Cham H. Investigating factorial invariance in longitudinal data. *Handbook of developmental research methods*. New York, NY, US: The Guilford Press (2012). p. 109-26.

451. Enders CK. The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Educ Psychol Meas* (2001) 61(5):713-40. doi: Doi 10.1177/0013164401615001. PubMed PMID: WOS:000171048800001.

452. Peeters M, Zondervan-Zwijnenburg M, Vink G, van de Schoot R. How to handle missing data: A comparison of different approaches. *Eur J Dev Psychol* (2015) 12(4):377-94. doi: 10.1080/17405629.2015.1049526. PubMed PMID: WOS:000359482500001.

453. Jackson H, Engelman M, Bandeen-Roche K. Robust Respondents and Lost Limitations: The Implications of Nonrandom Missingness for the Estimation of Health Trajectories. *J Aging Health* (2019) 31(4):685-708. doi: 10.1177/0898264317747079. PubMed PMID: WOS:000461623000008.

454. Muthen B, Asparouhov T, Hunter AM, Leuchter AF. Growth Modeling With Nonignorable Dropout: Alternative Analyses of the STAR*D Antidepressant Trial. *Psychological Methods* (2011) 16(1):17-33. doi: 10.1037/a0022634. PubMed PMID: WOS:000288300800002.

455. Enders CK. Missing not at random models for latent growth curve analyses. *Psychological Methods* (2011) 16(1):1-16. doi: 10.1037/a0022640.

456. Moore CM, MaWhinney S, Forster JE, Carlson NE, Allshouse A, Wang XS, et al. Accounting for dropout reason in longitudinal studies with nonignorable dropout. *Stat Methods Med Res* (2017) 26(4):1854-66. doi: 10.1177/0962280215590432. PubMed PMID: WOS:000407924800019.

457. Mehta PD, West SG. Putting the individual back into individual growth curves. *Psychological Methods* (2000) 5(1):23-43. doi: 10.1037//1082-989x.5.1.23. PubMed PMID: WOS:000088349900002.

458. Akaike H, editor. *Information theory and an extension of the maximum likelihood principle*. Budapest, Hungary: Akadémiai Kiadó (1973).

459. Schwarz G. Estimating the Dimension of a Model. *The Annals of Statistics* (1978) 6(2):461-4.

460. Sclove SL. Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika* (1987) 52(3):333-43. doi: 10.1007/BF02294360.

461. Ramaswamy V, Desarbo WS, Reibstein DJ, Robinson WT. An Empirical Pooling Approach for Estimating Marketing Mix Elasticities with Pims Data. *Market Sci* (1993) 12(1):103-24. doi: DOI 10.1287/mksc.12.1.103. PubMed PMID: WOS:A1993KY95900006.

462. Celeux G, Soromenho G. An entropy criterion for assessing the number of clusters in a mixture model. *J Classif* (1996) 13(2):195-212. doi: Doi 10.1007/Bf01246098. PubMed PMID: WOS:A1996VP96000001.

463. Muthén BO. What is a good value of entropy? (2007) [cited 2020 November 10]. Available from: http://www.statmodel.com/discussion/messages/13/2562.html?1237580237.

464. Hipp JR, Bauer DJ. Local solutions in the estimation of growth mixture models. *Psychological Methods* (2006) 11(1):36-53. doi: 10.1037/1082-989x.11.1.36. PubMed PMID: WOS:000236745400003.

465. Qizilbash N, Gregson J, Johnson ME, Pearce N, Douglas I, Wing K, et al. BMI and risk of dementia in two million people over two decades: a retrospective cohort study. *Lancet Diabetes Endo* (2015) 3(6):431-6. doi: 10.1016/S2213-8587(15)00033-9. PubMed PMID: WOS:000354880300020.

466. Nguyen JCD, Killcross AS, Jenkins TA. Obesity and cognitive decline: role of inflammation and vascular changes. *Frontiers in Neuroscience* (2014) 8. doi: ARTN 375; 10.3389/fnins.2014.00375. PubMed PMID: WOS:000346564500001.

467. Prickett C, Brennan L, Stolwyk R. Examining the relationship between obesity and cognitive function: A systematic literature review. *Obes Res Clin Pract* (2015) 9(2):93-113. doi: 10.1016/j.orcp.2014.05.001. PubMed PMID: WOS:000353068500001.

468. Rubin DB. *Multiple imputation for nonresponse in surveys*. New York ;: Wiley (1987).

469. Schafer JL. *Analysis of Incomplete Multivariate Data (1st ed.)*. New York: Chapman and Hall/CRC (1997).

470. Wayman JC. Multiple imputation for missing data: What is it and how can I use it? *Annual Meeting of the American Educational Research Association, Chicago, IL* (2003):2-16.

471. Nakai M, Ke W. Review of the Methods for Handling Missing Data in Longitudinal Data Analysis. *International Journal of Mathematical Analysis* (2011) 5(1):1-13.

472. Kelly-Hayes M, Wolf PA, Kase CS, Gresham GE, Kannel WB, D'Agostino RB. Time Course of Functional Recovery After Stroke: The Framingham Study. *Journal of Neurologic Rehabilitation* (1989) 3(2):65-70. doi: 10.1177/136140968900300202.

473. Tilling K, Sterne JA, Rudd AG, Glass TA, Wityk RJ, Wolfe CD. A new method for predicting recovery after stroke. *Stroke* (2001) 32(12):2867-73. Epub 2001/12/12. doi: 10.1161/hs1201.099413. PubMed PMID: 11739989.

474. Demain S, Wiles R, Roberts L, McPherson K. Recovery plateau following stroke: fact or fiction? *Disabil Rehabil* (2006) 28(13-14):815-21. Epub 2006/06/17. doi: 10.1080/09638280500534796. PubMed PMID: 16777768.

475. Hankey GJ, Spiesser J, Hakimi Z, Bego G, Carita P, Gabriel S. Rate, degree, and predictors of recovery from disability following ischemic stroke. *Neurology* (2007) 68(19):1583-7. Epub 2007/05/09. doi: 10.1212/01.wnl.0000260967.77422.97. PubMed PMID: 17485645.

476. Sprigg N, Gray LJ, Bath PMW, Lindenstrom E, Boysen G, De Deyn PP, et al. Stroke severity, early recovery and outcome are each related with clinical classification of stroke: Data from the 'Tinzaparin in Acute Ischaemic Stroke Trial' (TAIST). *Journal of the Neurological Sciences* (2007) 254(1-2):54-9. doi: 10.1016/j.jns.2006.12.016. PubMed PMID: WOS:000245391200011.

477. Jorgensen HS, Nakayama H, Raaschou HO, Olsen TS. Stroke. Neurologic and functional recovery the Copenhagen Stroke Study. *Phys Med Rehabil Clin N Am* (1999) 10(4):887-906. Epub 1999/11/26. PubMed PMID: 10573714.

478. Neiman J, Haapaniemi HM, Hillbom M. Neurological complications of drug abuse: pathophysiological mechanisms. *Eur J Neurol* (2000) 7(6):595-606. Epub 2001/01/03. doi: 10.1046/j.1468-1331.2000.00045.x. PubMed PMID: 11136345.

479. Fernández-Serrano MJ, Pérez-García M, Verdejo-García A. What are the specific vs. generalized effects of drugs of abuse on neuropsychological performance? *Neurosci Biobehav Rev* (2011) 35(3):377-406. Epub 2010/05/11. doi: 10.1016/j.neubiorev.2010.04.008. PubMed PMID: 20451551.

480. Schulte MHJ, Cousijn J, den Uyl TE, Goudriaan AE, van den Brink W, Veltman DJ, et al. Recovery of neurocognitive functions following sustained abstinence after substance dependence and implications for treatment. *Clinical Psychology Review* (2014) 34(7):531-50. doi: 10.1016/j.cpr.2014.08.002. PubMed PMID: WOS:000344445600002.

481. Rehm J, Hasan OSM, Black SE, Shield KD, Schwarzinger M. Alcohol use and dementia: a systematic scoping review. *Alzheimers Research & Therapy* (2019) 11. doi: ARTN 1; 10.1186/s13195-018-0453-0. PubMed PMID: WOS:000454955400003.

482. Brust JCM. Strokes and Substance Abuse. In: Biller J, Caplan L, editors. *Uncommon Causes of Stroke*. 3 ed. Cambridge: Cambridge University Press (2018). p. 429-36.

483. Moos RH, Brennan PL, Mertens JR. Mortality-Rates and Predictors of Mortality among Late Middle-Aged and Older Substance-Abuse Patients. *Alcoholism-Clinical and Experimental Research* (1994) 18(1):187-95. doi: DOI 10.1111/j.1530-0277.1994.tb00902.x. PubMed PMID: WOS:A1994MZ33500031.

484. Roerecke M, Rehm J. Cause-specific mortality risk in alcohol use disorder treatment patients: a systematic review and meta-analysis. *International Journal of Epidemiology* (2014) 43(3):906-19. doi: 10.1093/ije/dyu018. PubMed PMID: WOS:000338127000035.

485. Bill O, Zufferey P, Faouzi M, Michel P. Severe stroke: patient profile and predictors of favorable outcome. *J Thromb Haemost* (2013) 11(1):92-9. Epub 2012/11/13. doi: 10.1111/jth.12066. PubMed PMID: 23140236.

486. Treadwell SD, Thanvi B. Malignant middle cerebral artery (MCA) infarction: pathophysiology, diagnosis and management. *Postgraduate Medical Journal* (2010) 86(1014):235-42. doi: 10.1136/pgmj.2009.094292. PubMed PMID: WOS:000276081500008.

487. Balami JS, Chen RL, Grunwald IQ, Buchan AM. Neurological complications of acute ischaemic stroke. *Lancet Neurol* (2011) 10(4):357-71. doi: 10.1016/S1474-4422(10)70313-6. PubMed PMID: WOS:000289185000019.

488. Stern Y. What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society* (2002) 8(3):448-60. doi: 10.1017/S1355617702813248. PubMed PMID: WOS:000174602600010.

489. National Institute for Health and Care Excellence (NICE). Cardiovascular disease: risk assessment and reduction, including lipid modification. Cinical guideline [CG181]. (2014) [updated 27 Sep 2016; cited 2021 February 11]. Available from: https://www.nice.org.uk/guidance/cg181.

490. Lanza ST, Rhoades BL. Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment. *Prev Sci* (2013) 14(2):157-68. Epub 2011/02/15. doi: 10.1007/s11121-011-0201-1. PubMed PMID: 21318625.

491. Nys GM, Van Zandvoort MJ, De Kort PL, Jansen BP, Van der Worp HB, Kappelle LJ, et al. Domain-specific cognitive recovery after first-ever stroke: a follow-up study of 111 cases. *J Int Neuropsychol Soc* (2005) 11(7):795-806. Epub 2006/03/08. doi: 10.1017/s1355617705050952. PubMed PMID: 16519259.

492. Klimkowicz-Mrowiec A, Dziedzic T, Slowik A, Szczudlik A. Predictors of poststroke dementia: results of a hospital-based study in poland. *Dement Geriatr Cogn Disord* (2006) 21(5-6):328-34. Epub 2006/03/02. doi: 10.1159/000091788. PubMed PMID: 16508296.

493. Bell ML, Fairclough DL. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. *Stat Methods Med Res* (2014) 23(5):440-59. Epub 2013/02/22. doi: 10.1177/0962280213476378. PubMed PMID: 23427225.

494. Mueller RO, Hancock GR. Best Practices in Quantitative Methods. (2008 2021/02/15). [Internet]. Thousand Oaks, California: SAGE Publications, Inc. Available from: https://methods.sagepub.com/book/best-practices-in-quantitative-methods.

495. Schumacker RE, Lomax RG. *A beginner's guide to structural equation modeling*. Fourth edition. ed. New York, NY: Routledge (2016).

496. Appelros P, Nydevik I, Seiger A, Terent A. Predictors of severe stroke - Influence of preexisting dementia and cardiac disorders. *Stroke* (2002) 33(10):2357-62. doi: 10.1161/01.Str.0000030318.99727.Fa. PubMed PMID: WOS:000178515800011.

497. Saposnik G, Cote R, Rochon PA, Mamdani M, Liu Y, Raptis S, et al. Care and outcomes in patients with ischemic stroke with and without preexisting dementia. *Neurology* (2011) 77(18):1664-73. doi: 10.1212/WNL.0b013e31823648f1. PubMed PMID: WOS:000296590800007.

498. Holt-Lunstad J, Smith TB, Layton JB. Social Relationships and Mortality Risk: A Meta-analytic Review. *Plos Medicine* (2010) 7(7). doi: ARTN e1000316; 10.1371/journal.pmed.1000316. PubMed PMID: WOS:000280557200015.

499. Gallant MP. The influence of social support on chronic illness self-management: a review and directions for research. *Health Educ Behav* (2003) 30(2):170-95. Epub 2003/04/16. doi: 10.1177/1090198102251030. PubMed PMID: 12693522.

500. DiMatteo MR. Social support and patient adherence to medical treatment: a meta-analysis. *Health Psychol* (2004) 23(2):207-18. Epub 2004/03/11. doi: 10.1037/0278-6133.23.2.207. PubMed PMID: 15008666.

501. Essery R, Geraghty AW, Kirby S, Yardley L. Predictors of adherence to home-based physical therapies: a systematic review. *Disabil Rehabil* (2017) 39(6):519-34. Epub 2016/04/22. doi: 10.3109/09638288.2016.1153160. PubMed PMID: 27097761.

502. Corley J, Cox SR, Deary IJ. Healthy cognitive ageing in the Lothian Birth Cohort studies: marginal gains not magic bullet. *Psychol Med* (2018) 48(2):187-207. Epub 2017/06/10. doi: 10.1017/S0033291717001489. PubMed PMID: 28595670.

503. Stern Y. Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurol* (2012) 11(11):1006-12. Epub 2012/10/20. doi: 10.1016/S1474-4422(12)70191-6. PubMed PMID: 23079557.

504. Chapko D, McCormack R, Black C, Staff R, Murray A. Life-course determinants of cognitive reserve (CR) in cognitive aging and dementia - a systematic literature review. *Aging Ment Health* (2018) 22(8):915-26. Epub 2017/07/14. doi: 10.1080/13607863.2017.1348471. PubMed PMID: 28703027.

505. Carin-Levy G, Nicol K, van Wijck F, Mead GE. Delirium in acute stroke: A survey of screening and diagnostic practice in Scotland. *International Journal of Stroke* (2013) 8:74-. PubMed PMID: WOS:000330526900186.

506. Siddiqi N, Harrison JK, Clegg A, Teale EA, Young J, Taylor J, et al. Interventions for preventing delirium in hospitalised non-ICU patients. *The Cochrane database of systematic reviews* (2016) 3:CD005563. Epub 2016/03/12. doi: 10.1002/14651858.CD005563.pub3. PubMed PMID: 26967259.

507. Song J, Lee M, Jung D. The Effects of Delirium Prevention Guidelines on Elderly Stroke Patients. *Clin Nurs Res* (2018) 27(8):967-83. Epub 2017/07/25. doi: 10.1177/1054773817721400. PubMed PMID: 28737064.

508. Kent DM, Rothwell PM, Ioannidis JP, Altman DG, Hayward RA. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal. *Trials* (2010) 11:85. Epub 2010/08/14. doi: 10.1186/1745-6215-11-85. PubMed PMID: 20704705.

509. Kent DM, Hayward RA. Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification. *JAMA* (2007) 298(10):1209-12. Epub 2007/09/13. doi: 10.1001/jama.298.10.1209. PubMed PMID: 17848656.

510. Hayward RA, Kent DM, Vijan S, Hofer TP. Multivariable risk prediction can greatly enhance the statistical power of clinical trial subgroup analysis. *BMC medical research methodology* (2006) 6(1):1-11.

511. Collins R, Silarova B, Clare L. Dementia Primary Prevention Policies and Strategies and Their Local Implementation: A Scoping Review Using England as a Case Study. *J Alzheimers Dis* (2019) 70:S303-S18. doi: 10.3233/Jad-180608. PubMed PMID: WOS:000481670800023.

512. Shah H, Albanese E, Duggan C, Rudan I, Langa KM, Carrillo MC, et al. Research priorities to reduce the global burden of dementia by 2025. *Lancet Neurol* (2016) 15(12):1283-92. PubMed PMID: WOS:000386315700023.

513. Kivipelto M, Mangialasche F, Ngandu T. Can lifestyle changes prevent cognitive impairment? *Lancet Neurol* (2017) 16(5):338-9. doi: 10.1016/S1474-4422(17)30080-7. PubMed PMID: WOS:000398782100005.

514. Richard E, Van den Heuvel E, Moll van Charante EP, Achthoven L, Vermeulen M, Bindels PJ, et al. Prevention of dementia by intensive vascular care (PreDIVA): a cluster-randomized trial in progress. *Alzheimer Dis Assoc Disord* (2009) 23(3):198-204. Epub 2009/10/09. doi: 10.1097/WAD.0b013e31819783a4. PubMed PMID: 19812459.

515. Kivipelto M, Solomon A, Ahtiluoto S, Ngandu T, Lehtisalo J, Antikainen R, et al. The Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER): study design and progress. *Alzheimer's & dementia : the journal of the Alzheimer's Association* (2013) 9(6):657-65. Epub 2013/01/22. doi: 10.1016/j.jalz.2012.09.012. PubMed PMID: 23332672.

516. Vellas B, Carrie I, Gillette-Guyonnet S, Touchon J, Dantoine T, Dartigues JF, et al. Mapt Study: A Multidomain Approach for Preventing Alzheimer's Disease: Design and Baseline Data. *J Prev Alzheimers Dis* (2014) 1(1):13-22. Epub 2014/06/01. PubMed PMID: 26594639.

517. Brainin M, Matz K, Nemec M, Teuschl Y, Dachenhausen A, Asenbaum-Nan S, et al. Prevention of poststroke cognitive decline: ASPIS--a multicenter, randomized, observer-blind, parallel group clinical trial to evaluate multiple lifestyle interventions--study design and baseline characteristics. *Int J Stroke* (2015) 10(4):627-35. Epub 2013/11/12. doi: 10.1111/ijs.12188. PubMed PMID: 24206541.

518. Lennon S, McKenna S, Jones F. Self-management programmes for people post stroke: a systematic review. *Clin Rehabil* (2013) 27(10):867-78. Epub 2013/04/02. doi: 10.1177/0269215513481045. PubMed PMID: 23543340.

519. Sadler E, Wolfe CD, Jones F, McKevitt C. Exploring stroke survivors' and physiotherapists' views of self-management after stroke: a qualitative study in the UK. *BMJ Open* (2017) 7(3):e011631. Epub 2017/03/12. doi: 10.1136/bmjopen-2016-011631. PubMed PMID: 28283483.

520. Comas-Herrera A, Wittenberg R, Pickard L, Knapp M. Cognitive impairment in older people: future demand for long-term care services and the associated costs. *Int J Geriatr Psychiatry* (2007) 22(10):1037-45. Epub 2007/07/03. doi: 10.1002/gps.1830. PubMed PMID: 17603823.

521. Leslie DL, Inouye SK. The importance of delirium: economic and societal costs. *J Am Geriatr Soc* (2011) 59 Suppl 2:S241-3. Epub 2011/12/07. doi: 10.1111/j.1532-5415.2011.03671.x. PubMed PMID: 22091567.

522. Wu YT, Fratiglioni L, Matthews FE, Lobo A, Breteler MM, Skoog I, et al. Dementia in western Europe: epidemiological evidence and implications for policy making. *Lancet Neurol* (2016) 15(1):116-24. Epub 2015/08/25. doi: 10.1016/S1474-4422(15)00092-7. PubMed PMID: 26300044.

523. Public Health England. *From evidence into action: Opportunities to protect and improve the nation's health*. London, UK: PHE Publications (2014).

524. Bath PM, Krishnan K. Interventions for deliberately altering blood pressure in acute stroke. *The Cochrane database of systematic reviews* (2014) (10):CD000039. Epub 2014/10/30. doi: 10.1002/14651858.CD000039.pub3. PubMed PMID: 25353321.

525. Wang H, Tang Y, Rong X, Li H, Pan R, Wang Y, et al. Effects of early blood pressure lowering on early and long-term outcomes after acute stroke: an updated meta-analysis. *PLoS One* (2014) 9(5):e97917. Epub 2014/05/24. doi: 10.1371/journal.pone.0097917. PubMed PMID: 24853087.

526. Smith AD, Mangialasche F, Kivipelto M. Dementia research priorities-2. *Lancet Neurol* (2017) 16(3):181-2. Epub 2017/02/24. doi: 10.1016/S1474-4422(16)30401-X. PubMed PMID: 28229885.